

nature

27 October 1977

How much further can the pendulum swing?

IN THE past few years the tide of inflation has run so steadily in Britain, and the economy has been so sluggish, that we have rather lost sight of former landmarks and tended to measure performance mainly on the outcome of individual battles and on our ability to survive from year to year. Viewed in a longer time perspective, however, how has science fared in these years? In the Third Leverhulme Lecture in Liverpool last week, the biochemist Hans Kornberg pulled together a number of scattered statistics to show that although we may have survived the battles in some way or another, we are still in danger of losing the war. These matters are of particular interest now that North Sea oil is bringing some relief to a beleaguered economy and everyone who has been asked to make sacrifices is wondering when the sacrificing is going to stop.

Most striking of Kornberg's examples is the way in which expenditure on civil research and development in five European countries moved in the years 1969–74. West Germany, France and the United Kingdom spent respectively 1.5, 1.2 and 0.9 billion (10⁹) Eurodollars in 1969; by 1974 Germany was up to 3.6 billion, France to 2.0 billion . . . and the UK almost static at 1.0 billion. Expenditures in the Netherlands and Belgium, previously down at the 0.2–0.3 billion level, are steadily rising.

The figures do, of course, need extensive qualification. No one who has travelled on the Continent will be unaware of the discrepancies in cost of living which exchange rates seem poorly to reflect. And there are complications between countries in the allowance that is made for social security payments. So a better measure might well be how many people are employed in research and development per thousand of population in each country. But the general shape of the conclusion would be the same: that in Europe, Britain is lagging very seriously in its investment in civil R&D, though still spending heavily in the defence field in its urge to keep in touch with the United States.

Even within the term 'civil R&D', however, there needs to be some attention to detail. In recent years

the Rothschild reorganisation of government R&D and the more recent enthusiasm for engineering has ensured a moderate degree of health for the applied side of things. But with a static total budget this has inevitably meant harder times for the more basic side. In many ways the approach to these harder times has been too subtle to cause individual comment. One such way is the upwards creep of costs, necessary just to stand still. As staff become more senior, they are entitled to higher wages; as research proceeds, more complicated equipment is generally needed. Such costs outstrip cost-of-living scales normally used to adjust authorisations from year to year. As Kornberg puts it, "The pendulum has now swung too far [from the more affluent days of the 1960s] and its action must be reversed if it is not to cause possibly irreversible damage".

Few scientists would need to be reminded of one perfectly valid argument for supporting basic research—that you can never tell what it will turn up, and the occasional discovery, be it Newton's Laws or Mendel's peas or whatever, is infinitely more valuable in the long run than any amount of basic research. But there are two more arguments which should be deployed more frequently. They are, first, that support of basic research is support of a cultural activity, which government ought to have a genuine pride in fostering; and, second, that basic research and indeed basic researchers recognise no national frontiers, and so if support wanes in one country whilst waxing elsewhere British researchers will first have to depend more and more on the kindness of their foreign colleagues and ultimately will move themselves, or recommend their students to move to more fertile pastures.

The erosion is slow and there is no psychologically 'right' time at which to start protesting; besides, scientists recognise as well as anyone that there have been some genuine national needs for economy. But maybe Kornberg's speech, together with the Secretary of State for Education and Science's call a few months ago for a science lobby, will start something. □



The Nobel prizes (1): Physics

Rewarding solid state

Sam Edwards on Nevill Mott,
Philip Anderson and John van Vleck

IT is a pleasure to contribute a note on these awards, as I have had two of the recipients, Nevill Mott and Phil Anderson, as colleagues and have worked with them, and have attended the inspiring lectures of the third, John van Vleck, when a graduate student at Harvard.

Nevill Mott started research in physics when quantum mechanics was recently invented. His first great contribution was to show that quantum mechanics affected the most basic of scattering experiments, that of α -particle-helium collisions. Mott relates that, starting his first research appointment, he went to explain to Rutherford that his famous scattering formula needed amendment and that Chadwick, who came along, had found it was so. Rutherford listened with interest but without comment until Mott rose to leave, when the great man called after him, "If you think of anything else like this, please come and tell me".

This interest of Mott extended to atomic collisions and led to the collaboration with Sir Harrie Massey in their famous text on collision theory; it extended into nuclear scattering and very notably into the theory of internal conversion of gamma rays. His interests in the 1930s, after a move to Bristol, turned to the solid state, and there were many notable contributions to the theory of semiconductors, for example, to the understanding of rectification. This latter problem led him naturally to the study of oxides where he realised that the quantum theory of metals, semiconductors, and insulators, as developed by physicists in the 1930s, seemed quite inadequate to explain why under certain conditions oxides were conductors, whilst in other conditions they were insulators. People with a chemistry background found these facts explainable on structural grounds, but the basic interpretation of the solid state by quantum theory had now reached a level that anything other than a complete *ab initio* explanation of behaviour was inadequate.

It should be remarked that Mott had left Cambridge first for Manchester and then for Bristol where he played a major part in establishing the great eminence of the Bristol physics department, and made great contributions to the theory of dislocations, the theory of fracture and other aspects of the structure of solids, which has long

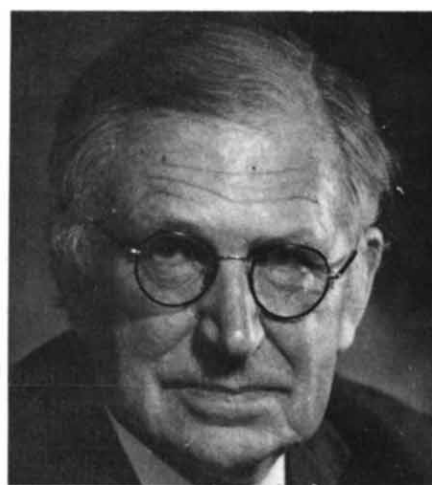
been a speciality of Bristol, and also to the theory of ionic crystals, a key to the photographic process which led to the famous book written with Gurney. He also wrote another famous book on metals in collaboration with Professor Jones, recently retired from Imperial College. But to return to the events which led to the Nobel Prize, Mott realised that external circumstances could lead a solid to make transitions between metal and insulator configurations, and the Mott transition was born. This study explained in a much deeper way than ever before why a material took up the metallic or non-metallic condition in the crystalline state, and is developed in detail in his book *Metal-Insulator Transitions*.

Back in Cambridge, he realised similar problems must exist in the amorphous state, which also shows metallic, insulator and semiconductor behaviour, and does so in a way which contradicts any straightforward approach. About this time Anderson published his paper 'Absence of diffusion in certain random lattices', which gave a rigorous basis to the phenomena of the localisation of electron wave functions by disorder. A lifetime of experience enabled Mott to see what the physical consequences of localisation would be, and a torrent of new physics resulted. Again this work formed the basis of a major book in collaboration with Dr E. A. Davis who has been Sir Nevill's experimental collaborator. Dr Davis tells me that no sooner does he prepare a new edition of the book than the number of papers in the literature justifies starting again. The phenomena flaring from the new theory include semiconductor switches and the prospect of much cheaper materials for, say, the conversion of sunlight to electricity, a field where Professor Spear has been quick to grasp the significance of the revolution and produce remarkable new amorphous materials.

The stature of Mott's work from the very beginning has been of the highest calibre, but it is perhaps most remarkable that the Nobel citation particularly referred to work done in his sixties, when most people are coasting to rest.

Anderson on magnetism

Phil Anderson started his research career studying aspects of magnetism with Van Vleck at Harvard, and early in his published papers is a notable study of the antiferromagnetic state, where he shows how quantum mechanics introduces great complexity into a problem which is classically so simple. His many early papers on magnetism led him to great facility with the application of quantum mechanical ideas in particular with spin systems into which he always tries to translate problems, since this way of thinking demands a quantal rather than classical intuition. It was natural therefore for him to think of the



Van Vleck, Anderson, Mott

diffusion of spins when he studied the problem of disorder in quantum mechanics in his famous 1958 paper referred to above. The same framework of thought produced impressive contributions to the theory of superconductivity which had been opened up around that time by the BCS theory.

When Anderson took a joint appointment in Cambridge and at Bell Telephones, one of the great fruits of his time in Cambridge was his interaction with Sir Brian Pippard and with Brian Josephson, when Josephson was developing new ideas on superconducting tunnelling which led to his Nobel prize. The other big interaction in Cambridge was with Mott and many other British physicists which has produced many new ideas both in Cambridge and elsewhere in Britain, notably with Professor Thouless in Birmingham.

Anderson also did basic work on helium 3 over this period, and even the application of solid state ideas to astrophysics, but it is his work on the electronic structure of disordered systems and on magnetism which has earned him the Nobel prize. Other notable work on glasses has been on the low temperature thermal properties, and on disordered magnets, the 'spin glasses' where it has been my privilege to collaborate with him.

I have noted Anderson's early and continued interest in magnetism and how it has influenced his other interests. That interest was kindled by J. H. Van Vleck, who brought

magnetism into the realm of quantum mechanics. It is astonishing to realise the rate of progress in understanding in the late 1920s and early 1930s when, compared to today, a small number of scientists revolutionised physics. Van Vleck published the first edition of his *Theory of Electric and Magnetic Susceptibilities* in 1932 and with this virtually the whole of the modern theory of magnetism was set up: not just solved because there are still deep and difficult problems remaining, but just how the classical theory could be absorbed into quantum theory, what the interactions had to be, how the Schrödinger equation was to be studied when crystal fields are present—it is all there, written out in a completeness and clarity which quite belie the date.

Subsequently Van Vleck developed and clarified many aspects of the electromagnetic nature of crystals, whilst also making major contributions to microwave and resonance spectroscopy, fields which have taken on new importance with the discovery of cosmic hydroxyl maser effects, and microwave communications. His work also extended to clarification of the nature of the chemical bond unifying the different valence bond and orbital pictures, but most importantly to the unification of the different models of magnetism. His Nobel award is something which his many advisers have been expecting for some years and gives just recognition for his achievement. □

The Nobel prizes (2): Chemistry

Thermodynamicist

J. S. Rowlinson on Ilya Prigogine

THE Swedish Academy of Sciences has awarded this year's Nobel prize for chemistry to Professor Ilya Prigogine of the Université Libre de Brussels. He was born in Moscow sixty years ago but went to Belgium at an early age, and has spent his whole career there, although for the past ten years he has also held an appointment at the Center for Statistical Mechanics and Thermodynamics at the University of Texas.

Brussels has had a long tradition of research in classical or equilibrium thermodynamics, and although Prigogine has contributed to this field he has done his most important work in two other branches, statistical thermodynamics, which is based on molecular descriptions, and non-equilibrium thermodynamics. He started in the first with some war-time papers which developed the cell theories of liquid structure put forward in the 1930s by Lennard-Jones, and developed them to cover liquid mixtures also. Such theories suppose that a molecule spends most of its time in one of a lattice of cells to which it is confined by its neighbours. To be applicable to liquid mixtures such theories require that each species can be accommodated equally well in each cell.

This work, and parallel work in England and America in the 1950s, was a real advance, but was restricted in its useful application to mixtures of molecules of the same size. His book *Molecular Theory of Solutions* (1957) was influential in its day but is not much used now. The theory of liquids has developed along other lines which eschew the restrictive concepts of lattices or cells.

Meanwhile Prigogine was putting more of his effort into non-equilibrium thermodynamics, a subject which has had a long but not always very productive history. It started in the last century with the study of apparently simple transport processes such as the flow of electric current caused by a temperature difference between two junctions of different metals (Seebeck effect), or the inverse, the evolu-

tion of heat produced by an electric current across a metallic junction (Peltier effect).

Similar coupled flows of matter and energy occur when two fluids mix by diffusion, or when chemical reactions occur in inhomogeneous systems. Kelvin put forward quasi-thermodynamic explanations of some of these phenomena but it was not until 1931 that Onsager brought order to the whole subject by describing exactly how such forces and flows are coupled—his famous 'reciprocal relations' for which he was awarded the Nobel Prize for Chemistry in 1968. An interesting consequence of Onsager's work was Prigogine's discovery in 1945 that such systems are characterised by minima in the rates of entropy production.

These results are, however, limited to systems which are very close to equilibrium so that the flows are always proportional to the gradients of potential which generate them, such as temperature, or electric or chemical potential. They are, therefore, restricted in their applications, particularly if the systems are chemically reactive. One critic of this stage of development of the subject, believing that the useful results were not commensurate with the effort expended, said in 1966 that we should "ask not what we can do for irreversible thermodynamics; ask what irreversible thermodynamics can do for us". It was from this impasse that Prigogine sought to escape.

Within the regime of linear laws he undertook the development of molecular or statistical theories of non-equilibrium thermodynamics. Until the 1950s such theories had been confined to dilute gases and had reached their climax with the work of S. Chapman and D. Enskog during the First World War. The methods they used, however, and even the name of the subject—the kinetic theory of gases—had little in common with equilibrium



Prigogine

statistical thermodynamics. Prigogine, Résibois, and others of the Brussels school have played a leading part in developing kinetic theories that are applicable to all fluids. One of the principal tools of this research has been to exploit the analogies between the Liouville equation, which describes how a molecular system evolves, and the time-dependent Schrödinger equation of quantum mechanics.

A second line of attack has been to push the macroscopic treatment of Onsager beyond the narrow range of situations in which flows depend linearly on forces. If this could be done then a whole range of new phenomena would come within the scope of non-equilibrium thermodynamics, such as the turbulent flow and convective instabilities of hydrodynamics, and chemically reactive systems that oscillate in space or time. However a macroscopic description implies that we can define such statistical properties as pressure, temperature and entropy even if the system is far from equilibrium. Prigogine showed that such an assumption is a sufficient condition for the use of thermodynamic methods and justified it in certain cases, for example, by showing that even in fast gas reactions the molecular distribution of velocities shows little departure from a local Maxwellian distribution.

The behaviour of fluctuations of these local properties distinguishes stable from unstable systems. Spontaneous fluctuations are always occurring but, except at a critical point, are usually unimportant. In a system with weak potential gradients each fluctuation quickly regresses to the steady-state behaviour, but when gradients are large then the fluctuations can grow until a new pattern of flow is established, for example the Bénard cells of rotating elements of fluid in a liquid heated strongly from below. Prigogine has sought to marry the classical theories of

stability of structures with a generalised thermodynamics which would encompass such dissipative systems. He does this by introducing a 'local potential', which is a generalisation of the more familiar chemical potential, and then by using a variational principle to calculate the flows. The boundary between the stable or steady states and the dissipative structures is described by an 'excess entropy production' which is positive for the former and negative for the latter.

These new methods encompass many of the traditional ones used in discussing hydrodynamic instabilities, but their more interesting applications lie in the field of chemical reactions. This is, first, because reaction rates are, in practice, never linear in differences of chemical potential, and secondly because of the biological occurrence of systems in which complex chemical reactions are coupled to flows of matter and energy. He and his colleague, G. Nicolis, have used these methods to investigate the conditions under which reaction systems can oscillate in space and time with the aim of developing a thermodynamics that is applicable to biochemical systems.

Whatever new results emerge it is clear that Prigogine, both by his writings and by his enthusiastic advocacy, has changed the ways in which we can discuss a variety of difficult problems. He is, above all, a man of original ideas. In his book (with P. Glansdorff) *Thermodynamic Theory of Structure, Stability and Fluctuations* (1971), he quotes Bergson's saying of 1907 that the second law of thermodynamics is the most metaphysical of all laws of nature. He adds, "Whether a compliment or a criticism, this applies also to the 'generalised thermodynamics' we develop in this monograph".

□

The Nobel prizes (3): Physiology and Medicine

Starting a revolution

Jesse Roth and J. E. Rall on Rosalyn Yalow

DR Rosalyn Yalow was awarded a Nobel prize for medicine and physiology in 1977 for the development of the radioimmunoassay. The introduction of the radioimmunoassay for plasma insulin in humans (R. S. Yalow and S. A. Berson, *Nature* **184**, 1648; 1959; *J. Clin. Invest.* **39**, 1157; 1960) started a revolution in endocrinology which spread from insulin to other peptide hormones including substances previously thought not to be antigenic, to steroid and thyroid hormones and to the cyclic nucleotides. Like other great revolutions it spread rapidly to encompass other fields including clinical pharmacology, enzymology, oncology, virology, immunology and haematology. The recent explosion of studies of cellular receptors for hormones, neurotransmitters, lipoproteins and drugs has its technical and intellectual roots in the radioimmunoassay.

Missing from the Nobel citation is the name of Dr Solomon A. Berson who died in 1972. Nobel prizes are not awarded posthumously, but historians and scientists everywhere recall the hundreds of joint publications of Yalow and Berson in the two decades that flank the publication of the seminal papers that introduced the radioimmunoassay.

To appreciate the revolution that the radioimmunoassay wrought it is necessary to recall the incredible state of endocrinology in 1960. The peptide hormones, which constitute about 80% of all hormones, could not, with rare exception, be measured in blood or other biological fluids. If one wished to study the dynamics of insulin *in vivo*, one could only guess from measurements of the blood glucose, or

some other very distant component in the complex of reactions that regulate the blood glucose. Biological assays were available to measure the high concentrations of hormones in extracts of glands but these assays were orders of magnitude too insensitive and lacked the necessary specificity for measurements of hormones in blood.

How was it possible to measure peptide hormones at concentrations of 10^{-10} M in a sea of contaminating proteins at 10^{-3} M, when these hormones were virtually indistinguishable from thousands of other peptides and proteins in the blood? To solve this problem Berson and Yalow used antibodies to insulin, which they had recently characterised and had shown were capable of binding insulin but no other substances. Insulin that they labelled with radioactive iodine was reacted with antibody that had the desired specificity and the then unheard of affinity constant 10^{10} L/M. Conditions were set so that minute concentrations of unlabelled insulin competed with the labelled insulin for binding to the antibody. By measuring the ratio of radioactively-labelled insulin that was bound to antibodies to that which was free, the precise concentration of unlabelled insulin in the system could be determined.

This simple method of measurement of hormones in the blood was made possible by a series of conceptual leaps and technical innovations by Yalow and Berson in the preceding five years. They discovered that all insulin-treated patients generated antibodies to the hormone and recognised the problem of species specificity among insulins. They recognised the need to find antibodies with affinities so high as to be undreamed of by immunochemists of those days. They introduced methods to protect and purify proteins that were being labelled at high specific radioactivity. They devised a rapid precise method for separating antibody-bound hormone from free hormone and characterised in detail the kinetics and thermodynamics of insulin antibody interaction.

With precise and specific measurements for plasma hormone, it became clear that many of the earlier guesses were incorrect. For example, the radioimmunoassay proved that only a minority of diabetics have an absolute deficiency of insulin. Most diabetics have normal or supernormal amounts of insulin which is biologically intact but the cells of these diabetics are subnormally responsive to insulin. Yalow and Berson not only introduced the radioimmunoassay for insulin but applied the method to study the physiology of many other peptide hormones, including growth hormone, parathyroid hormone, adrenocorticotropin and gastrin. They also used this approach to devise a very sensitive method for detecting hepatitis virus and its antibodies in blood, which is widely used in blood banks in the United States and elsewhere. Other applications to problems of infectious diseases are now emerging. Although just beginning, the extension of radioimmunoassay to measure blood levels of drugs, including cardiac glycosides, anticonvulsants and antibiotics, has made it much more feasible for physicians to achieve therapeutic effects without sustaining serious untoward effects of these agents. The principle of competitive binding assays has been extended beyond antibodies to other binding substances of appropriate specificity and affinity.

In the light of these extraordinary scientific accomplishments, it is interesting to recall Dr Yalow's own modest personal and academic background. Born in 1921, she was raised in the South Bronx, the area that is now recognised as the most devastated in the urban United States. She was the first physics major at Hunter College, and although she graduated *magna cum laude* and *phi beta kappa*, her future was clouded by the fact that she was both Jewish and female, two traits which were not considered to favour success in physics in those days. In part because the military draft was depleting the ranks of eligible males, Dr Yalow was accepted by the University of Illinois, and after receiving her Ph.D in physics she returned to New York, where she worked until the end of the Second World War. In 1947, when applications of radioactive isotopes to clinical medicine were in their infancy, Dr Yalow joined the newly formed radioisotope unit of the

Veterans Administration Hospital in the Bronx where she has worked ever since.

Dr Berson, from a similar background, joined the radioisotope unit in 1950. His participation in the research was much less direct after 1968 when he assumed the chairmanship of the Department of Medicine of the Mount Sinai School of Medicine. In spite of his sudden death four years later, there has continued a flood of exciting new studies from the laboratory (renamed the Solomon A. Berson Research Laboratory) at the Bronx Veterans Administration Hospital under Dr Yalow's solo leadership. It has become clear that each peptide hormone in the blood is not a single substance but is rather a family of related peptides that include the active hormone, modifications of the active hormone, precursors and degradation products, as well as phylogenetically related peptides. Dr Yalow has been the major contributor to this area over the last decade and has extended these observations to numerous biological problems *in vivo*.

Her achievements are the subject of editorials in *The Daily News* and *The New York Post* as well as the subject of hundreds of sermons in churches and synagogues throughout the area. Dr Yalow is being viewed as an example of how talented and determined people of modest background can become world champions even in today's America. Women's groups all over North America have turned to her as a symbol. The Urban Crisis Task Force, which has its headquarters in the South Bronx, sees her as an ideal and has decided to name one of their new housing projects in her honour. □



Yalow

Inadvertent collaboration

George Fink on Roger Guillemin and Andrew Schally

DR Roger Guillemin of the Salk Institute and Dr Andrew Victor Schally of the Veterans Administration Hospital, New Orleans, were awarded the Nobel prize in physiology and medicine for isolating, characterising and synthesising three polypeptides which mediate the neural control of the anterior pituitary gland. The discoveries of Schally and Guillemin have already proved significant for clinical and basic medical science and are likely to offer new and safer methods for the control of population size.

The delight of neuroendocrinologists will only be dampened by the fact that the untimely death in November 1971 of Geoffrey Wingfield Harris, FRS, prevented him from sharing the fruits of his labours. As a medical student at Cambridge, Harris was the first (1937) to provide experimental proof for the then tentative view that the anterior pituitary gland was controlled by the central nervous system. The elegant studies carried out by Harris in the 1940s and early 1950s, alone and in collaboration with Dora Jacobsohn and the late John Green, established beyond doubt that this control was mediated by a neurohumoral mechanism involving the transport by hypophysial portal

vessel blood of chemical substances from the hypothalamus to the anterior pituitary.

The three polypeptides and their amino acid sequences are: thyrotrophin releasing factor (TRF) pyro Glu-His-Pro-NH₂; gonadotrophin releasing factor (GnRF) pyro Gly-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH₂; and somatostatin (or somatotrophin release inhibiting factor, SRIF) H-Ala-Gly-Cys-Lys-Asn-Phe-Phe-Trp-Lys-Thr-Phe-Thr-Ser-Cys-OH. Synthetic TRF and GnRF are now widely used in the investigation of thyroid dysfunction and infertility, respectively. Because TRF releases prolactin as well as thyrotrophin, the tripeptide is also used for investigating patients in whom prolactin secretion appears to be abnormal. Conceivably, antagonists of TRF may prove useful for controlling hyperprolactinaemia, a condition frequently associated with infertility in women, and for the control of hormone dependent tumours of the breast.

Superactive analogues of GnRF, made by substituting a D-amino acid (for example D-Trp, D-Ala or D-Leu) for Gly⁶ and an ethylamide group for Gly¹⁰, are being investigated mainly by Schally's group with a view to improving the precision of ovulation in women who prefer to or must use the rhythm method of fertility control. The parent decapeptide has already been used successfully for the treatment of certain types of infertility in men and women. Because GnRF will only discharge the amount of luteinising hormone (LH, the hormone which triggers ovulation) normally available for release, the risk in terms of pro-

ducing multiple pregnancy is negligible compared with that of administering exogenous gonadotrophins. Concern about the safety of the steroid contraceptive 'pill' has added impetus to the search for other chemical contraceptives.

As well as inhibiting the release of growth hormone, somatostatin inhibits the release of thyrotrophin, prolactin, insulin, glucagon and gastrin. In spite of these widespread effects, early trials suggested that the tetradecapeptide may prove useful for the control of certain aspects of diabetes mellitus. An inhibitory effect of somatostatin on platelet aggregation has, however, made it mandatory for clinicians to proceed with caution.

As expected, access to large amounts of the three peptides, which are relatively easily synthesised, has proved invaluable for investigating the physiology and biochemistry of the hypothalamic-pituitary system. Perhaps of greater importance, however, is the fact that the peptides have reawakened interest in extra-hypothalamic peptidergic neurons, a field which has largely remained dormant since the studies of Gaddum and Von Euler on substance P (1931). Using antisera to TRF, GnRF and somatostatin, immunoassay and immunohistochemical studies have shown that relatively large amounts of these peptides are present in areas outside the hypothalamus, including the cerebral cortex and spinal cord.

The small amount of polypeptide in the hypothalamus (nanogram quantities), the protected N and C terminals of TRF and GnRF, and the fact that the hypothalamus contains more pharmacologically active substances than any other tissue, made the isolation of the three peptides exceedingly difficult. Both Schally with Murray Saffran at McGill, and Guillemin in Houston, started their isolation studies on corticotrophin releasing factor (CRF). Subsequently, at Baylor University, they collaborated on this project, but, by the early 1960s, they realised that for various technical reasons, notably the lack of a specific assay for CRF, it would have been better to leave the investigation of this potentially important factor until last.

After Schally became chief of the Veterans Administration Endocrine and Polypeptide Laboratory in New Orleans (1963), the alarming rivalry which soon developed between the two Laureates provided the necessary stimulus for the completion of their Herculean task. Almost every result obtained by one was checked by the other, and, although separated more by passion than distance, much of the work proved inadvertently to be collaborative. Thus, for example, Guillemin and his co-workers were the first to show that hypothalamic extracts could release thyrotrophin (*C.R. Acad. Sci. Paris* **255**, 1018, 1962). Schally's group revealed that His, Glu and Pro were the aminoacids present in porcine TRF (*Biochem. Biophys. Res. Commun.* **25**, 165, 1966). On the basis of the activity of various synthetic tripeptide isomers of these acids, both groups reached the conclusion in 1969 that TRF was likely to be pyro Glu-His-Pro-NH₂.

Shortly afterwards, Guillemin reported that the structure of natural ovine TRF was identical to that of the active isomer as assessed by infrared, mass and nuclear-magnetic resonance spectrometry (*C.R. Acad. Sci. Paris* **269**, 226, 1969; *Nature* **226**, 321, 1970). Schally, found that natural porcine TRF was also identical to pyro Glu-His-Pro-NH₂ as assessed by identical R_f values in various chromatographic systems (*Biochem. Biophys. Res. Commun.* **37**, 705, 1969) and by spectral analysis (*Biochemistry* **9**, 1103, 1970).

Samuel (Don) McCann in 1960 and Geoffrey Harris in 1961, using respectively the ovarian ascorbic acid depletion method of Albert Parlow (an assay used extensively by Guillemin and Schally) and the rabbit ovulation test, reported that hypothalamic extracts released LH. After ten years of intense work carried out by several groups, Schally was able to announce in 1971 that his team had determined

the structure of porcine GnRF (*Biochem. Biophys. Res. Commun.* **43**, 393 and 1334). The amino-acid sequence was confirmed by Guillemin's group who used ovine tissues (*Proc. Natn. Acad. Sci.* **69**, 278, 1972). Both groups agreed that the decapeptide releases LH and follicle-stimulating hormone (FSH) (hence Schally's preferred term, LH-RH/FSH-RH). Earlier evidence for the existence of a separate releasing factor for FSH appeared to be due to the non-specific effects of polyamine contaminants. So far there is no convincing biochemical evidence for the existence of a FSH-RF distinct from the decapeptide, and this is disappointing since it makes remote the possibility of developing an acceptable male contraceptive 'pill'.

McCann and his co-workers (*Endocrinology* **83**, 783, 1968) first demonstrated the presence in fractions of sheep and rat hypothalamic extract of a growth hormone inhibitory factor and a growth hormone releasing factor. The structure of ovine growth hormone inhibitory factor (somatostatin) was reported by Guillemin's group in 1973 (*Science* **179**, 77). In spite of extensive studies the structure of growth hormone releasing factor remains unpublished.

Since the isolation of the three polypeptides, Schally's team have focused attention on the development of potent agonists and antagonists of GnRF, as well as on the isolation of prolactin-inhibitory, corticotrophin-releasing and growth-hormone-releasing factor. As well as developing analogues of the peptides, Guillemin has concentrated his efforts on the endorphins, his most significant recent contribution being the isolation of a pituitary protein (MW 31000) which appears to be a common precursor of β -endorphin and ACTH.

Apart from their expertise in chemistry and physiology, and their tenacity and perseverance, the success of the two Laureates depended on the recognition in the early sixties that the isolation of each hypothalamic factor required an all or none approach. They had the capacity to attract research funds amounting to many millions of dollars which enabled them to assemble large teams of competent chemists and biologists, and as well, purchase the many hundreds of thousands of hypothalamic fragments necessary for successful isolation.

Since the late 1950s, when he began to test ovine hypothalamic fragments for LH-releasing activity, Geoffrey Harris was determined to isolate LH-RF. He managed to secure only one senior chemist at a time: first Peter Fawcett at Mill Hill and then (1966) Harry Gregory at ICI. In the light of the resources available to the American teams, it is perhaps remarkable that, using only a few thousand fragments, Gregory had by 1970 isolated six of the aminoacids in GnRF and had shown that this factor had no free α amino groups (*Control of gonadal secretion*, eds. Baird and Strong, page 15). Is there a lesson here for the funding of British biomedical research?



Guillemin, Schally

JET

Permission to land

The delays experienced by the EEC's Joint European Torus (JET) fusion project are over. Chris Sherwell reports

A FULL two years after the matter came before it, the EEC Council of Research Ministers has agreed on the research laboratories at Culham near Oxford as the location for the next stage of the Community's fusion project, JET. The choice, made at a Council meeting on Tuesday, ends a controversial wrangle that has damaged the Community as well as the project. The decision also paves the way for the establishment of a JET Council, the release of funds to boost a stalling research effort and the appointment of a project head.

Political rather than scientific disputes have persistently hamstrung the £100-million project. Contracts for the team based at Culham have had to be extended at the last moment many times, and continuing uncertainties have driven some members away. Only last week Paul Rebut, the present project leader and a strong candidate to head the next phase, warned that rebuilding the team would take at least six months. Over the past 18 months JET has come up at five Research Ministers' meetings, two Foreign Ministers' meetings and one heads of government meeting; there have been six elections, and seven research ministers have changed; so have six foreign ministers and four heads of government.

The site of the project has sparked most controversy. National interest arguments prevented agreement on any of the four sites originally pro-

posed (Ispra in Italy, Cadarache in France, Garching in Germany and Culham in Britain), even when they were eventually whittled down to Garching and Culham. A widely expected decision failed to materialise at a Research Council meeting in March this year, and EEC heads of government passed the matter three months later to the Foreign Ministers. When they too failed to agree, everything seemed to hang on Anglo-German exchanges. But the Schleyer kidnapping intervened, no exchanges took place and JET was not discussed at a Foreign Ministers' meeting in September.

The breakthrough finally came last week when the West German Chancellor, Helmut Schmidt, and the UK Prime Minister, James Callaghan, met in the aftermath of the successful Mogadishu raid. Its euphoric effects, and Britain's supporting role in it, may have helped the two men submerge differences on matters like Community budget contributions. They agreed that a site for JET should be decided by a consensus—in other words, that neither country would veto whatever preference the majority expressed. This was conveyed to the Foreign Ministers' meeting in Luxembourg the same day, and though Germany did not waive her claim for Garching, the immediate presumption, based on past voting intentions, was that JET would go to Culham when the Research Ministers met in Luxembourg. In the event, five countries voted for Culham and two for Garching. One of the two abstaining countries went with the majority. Earlier fears of an Irish veto proved unfounded.

Reports that two JET projects were being discussed gained credence at a press conference in Brussels on Mon-

day given by Guido Brunner, the EEC's Energy and Research Commissioner. Culham and UK Foreign Office spokesmen last week discounted the idea, the former pointing out that even if this referred to the possibility of siting 'Super JET' or 'Jumbo JET' (as the next phase is known) in the country which failed to win JET, such a decision could not be taken bilaterally. The impression nevertheless remained on Monday that the Research Ministers would take a consensus view on locating the next phase at the losing site, and that ancillary fusion research might also be used as compensation.

The removal of other potential obstacles at a meeting of high-level officials a week earlier also assisted agreement on a JET site. Helpful concessions from France, which originally proposed a JET Council to manage the project and saw it become a stumbling block, produced agreement that the member states, along with other participating countries (Sweden and Switzerland), would each have two members on it, one a scientific expert. Voting will be weighted, 26 out of 37 constituting a majority. The Council, which will have a scientific council to consult at its convenience, will meet as and when necessary. Once it is established (initially in interim form) and funds become available, the senior scientific appointment for the project will be made. But that could prove as politically contentious as anything that has befallen JET so far.

The aim behind the £100-million project—often forgotten by the politicians, often simplified as "cheap, safe and abundant energy supplies"—is essentially to construct a third-generation tokamak device. The equivalent device in the USSR is known as T-20, in the USA as TFTR, and in Japan as T-60. □

DIOXIN

Informing on 2,4,5-T

Alastair Hay looks at two recent reports on the controversial herbicide 2,4,5-T

MANUFACTURERS and users of the herbicide 2,4,5-T are not letting attacks on it pass without comment, if two recent reports that attempt to vouch for its safety are anything to go by. The first, a Dow Chemical Company report to the US Environmental Protection Agency (EPA), insists that there is no "significant hazard" to humans asso-

ciated with its use. The second, from the UK Forestry Commission, claims that if foodstuffs or water supplies should become contaminated with 2,4,5-T, "serious consequences [for humans or animals] are unlikely" to occur.

The use of 2,4,5-T has aroused concern since the late 1960s when formulations of the herbicide in use in Vietnam were shown to be contaminated with 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (dioxin), which is highly toxic and a known teratogen. The public outcry,

together with scientific criticism of the use of 2,4,5-T, led the EPA to prohibit its use around homes, recreational sites, aquatic areas and on human food crops. In 1975 an EPA study reported that dioxin had been found in beef fat taken from cattle in areas where 2,4,5-T was known to have been used.

The Dow study was undertaken to assess these EPA findings. In a two-year study Dow scientists fed rats diets containing 0.1, 0.01 and 0.001 $\mu\text{g kg}^{-1}$ of dioxin a day. The animals receiving the highest dose experienced the known clinical and pathological symptoms of dioxin exposure—decreased weight gain, decreased red blood cell and

haemoglobin levels, elevation of some serum enzymes, and increased mortality. Discernible increases in the incidence of liver cancer and carcinomas of the lung, hard palate and tongue were also observed. But the report also speaks of a remarkable decrease in the "incidences of tumours of the pituitary, uterus, mammary glands, pancreas and adrenal gland" in these animals. In the rats receiving the lower doses of dioxin a lesser degree of toxicity was observed but there was no increase in tumours. On the basis of these results the report states that if the EPA measurements are used then the average person would need to consume fifty times his "body weight in beef every day to receive the daily dose level" of dioxin linked with cancer incidence in Dow's studies. Thus, in Dow's view it appears that 2,4,5-T is safe if used properly.

The same conclusion is drawn by the Forestry Commission report, 'The Safety of Herbicides 2,4-D and 2,4,5-T'. The author, D. J. Turner, describes many of the criticisms of the use of 2,4,5-T as unfounded and ill-informed. He points out that 2,4,5-T, still enjoying widespread use in Europe and South-East Asia, remains the most effective herbicide for defoliating rubber trees, which is the way to contain and isolate the South American

leaf blight disease *Dothidella ulei*. But he also says that "if reaction conditions [for converting tetrachlorobenzene to trichlorophenol—the precursor of 2,4,5-T] are properly regulated dioxins are not formed". Chemists knowledgeable in reaction mechanisms say this is not so, arguing that the reaction makes dioxin contamination unavoidable and that control of the reaction conditions will only limit the amount of dioxin produced, not eliminate it entirely. In their view the manufacturing stage presents greater risks of dioxin contamination than the use of the herbicide in silviculture.

In spite of the now considerable volume of literature relating to dioxin's effects, the report claims there is no evidence that the "defoliation spraying programmes in Vietnam had any direct effects on humans". This is flatly contradicted in the fifteenth report of the International Agency for Research on Cancer (IARC). In a chapter on chlorinated dibenzodioxins that cites the same National Academy of Sciences report referred to in the Forestry Commission publication, this states that the Montagnard people of South Vietnam were badly affected by exposure to the herbicide 'Agent Orange', a 50:50 mixture of the *n*-butyl esters of 2,4-D and 2,4,5-T,

containing up to 30 mg kg⁻¹ of dioxin. The consequences included diarrhoea, vomiting, skin rashes, fever, abdominal pain and the deaths of some children. The IARC publication, one of the most comprehensive reports yet produced on the dioxins, aimed to assess carcinogenicity of chlorinated dibenzodioxins, but concluded that this evaluation could not be made with the data available to it at the time of publication.

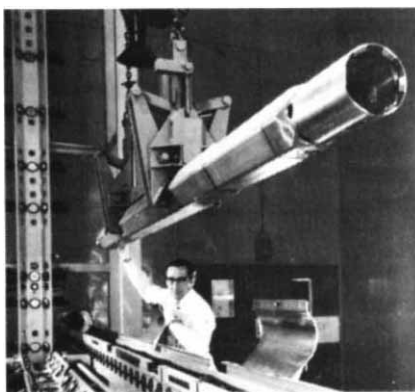
At Séveso in Italy—where the authorities are now satisfied that decontamination following last year's release of dioxin has made it safe for people to return to their homes (last week the first 24 families were allowed back)—135 children have chloracne as a result of dioxin exposure. A total of thirty-three adults but no children have neurological disorders, yet if (as some argue) chloracne is an obligate sign of dioxin poisoning, children might have been expected to have neurological problems. No eye damage has been reported in the 400 people examined by an oculist. Sixteen cases of children born with malformations were reported by August this year, of whom two died, but the diversity in the type of malformation and the fact that the number was no greater than the statistical norm suggests that dioxin was probably not the causative agent.

The development of a new chelating agent does not usually make headlines in the national press, but this week saw the widely publicised announcement of a chelating agent with two important differences. Its scientific interest lies in the fact that it can enter the cells of the body and will therefore be much more efficient at removing heavy metal contamination. The reason it hit the headlines is that it has been developed (at the Harwell laboratories of the UK National Radiological Protection Board) to remove plutonium from the body in the event of an accident.

One of the problems in devising effective treatments for the rare cases of accidental contamination with plutonium has been the difficulty of removing the plutonium which becomes locked up in the cells of the body, mostly in liver and bone, where over the years there is the risk that it might induce cancers. The only available treatment up to now aimed to inactivate the plutonium when it was freely accessible in the blood before it became deposited inside tissue cells. The chelating agent calcium diethylene - triamine - pentaacetic acid (CaDTPA) can pick up plutonium from the bloodstream and the resulting complex is excreted in the urine.

But CaDTPA has one great drawback—it cannot enter cells and so is only effective against plutonium circulating in the blood. In real life the

Fighting plutonium



Handling plutonium fuel assembly (BNFL)

BACKGROUNDER

contaminant may well be 'insoluble' plutonium, such as plutonium dioxide particles, which leach very slowly into the bloodstream from the lungs; in this case DTPA is not particularly efficient in reducing the total body burden.

The NRPB's new drug, which it calls Puchel, is a derivative of DTPA

with one important modification. Fatty acid chains have been added to the molecule enabling it to penetrate the lipid cell membrane and pick up plutonium from inside the cell. The complex is then excreted, mostly in the faeces. The NRPB is not yet divulging the structure of its new molecule, pending a possible patent application, but the secret apparently lies in the length of the fatty acid chains.

So far, of course, the drug has only been tried out in animals, where it has proved essentially non-toxic when given as the calcium salt. Injected intraperitoneally into hamsters it released most of the plutonium deposited in the liver after injection of a soluble plutonium salt, and given as a aerosol removed plutonium from the lung after inhalation of plutonium particles. It will also probably be effective in removing intracellular metal if given orally. Puchel has proved less effective at releasing plutonium from bone than from liver.

Puchel has been developed as a contribution towards reducing the risks of a nuclear accident, but if it proves itself it may be useful in cases of poisoning with non-radioactive heavy metals such as lead and mercury.

IN BRIEF

ISEE OK

The first stage of the joint European-American International Sun-Earth Explorer (ISEE) satellite mission was launched successfully at the weekend. Lift-off was postponed from 13 to 19 October and then suffered two further delays. This was to allow NASA to adjust the Thor Delta vehicle, launching the ISEE A and B satellites in tandem. The second part of the mission, ISEE C, will be launched in nine months' time.

Whale decision

A Federal Court of Appeal has rejected the Alaskan Eskimos' plea that the US Administration be required to file an objection against the International Whaling Commission's (IWC) decision

to bar the hunting of bowhead whales next year. The Court upheld the US decision not to object to the ban. It did so on the grounds that a US refusal to abide by IWC recommendations would give other countries licence to follow suit and would seriously threaten the bowhead with extinction, whose numbers are estimated at between 500 and 2,000.

Washington meeting

An idea to curb the spread of nuclear weapons by setting up an International Fuel Bank was suggested by President Carter last week at the opening of the 40-nation international conference on nuclear fuel disposal. The Bank would ensure the US and other countries of an adequate supply of uranium for

peaceful purposes. The US government also has a plan for storing all spent nuclear fuel produced by reactors both at home and abroad, which would further reduce the Administration's fears of unchecked proliferation.

Med meeting

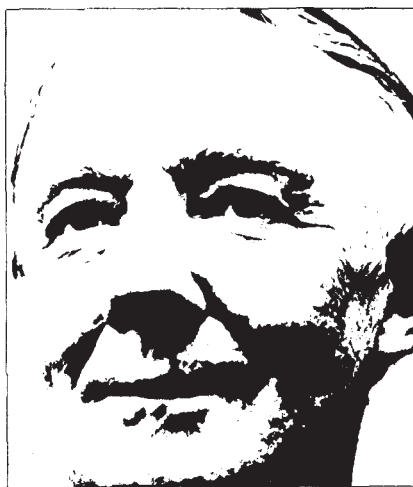
A five-day meeting of 13 Mediterranean states and the EEC to discuss the draft of a treaty to fight Mediterranean pollution, ended this week with agreement on limiting or banning the discharge of certain materials. The United Nations Environment Programme (UNEP) estimates that the cost of implementing the treaty, which will be submitted to Mediterranean states at a major conference next January, at \$5 billion.

CONSERVATIONISTS, anxious to preserve the native animals and plants of the countryside, generally assume that modern agricultural techniques are increasingly damaging to wildlife. The wholesale removal of hedges in the cereal growing areas of eastern England has, locally, left few wild habitats, spray drift from carelessly used weedkillers has denuded roadside verges of many attractive flowers, and some insecticides have slaughtered thousands of songbirds. However, these abuses are being generally controlled today, and in Britain there is growing cooperation between farmers, the makers of agricultural chemicals and conservationists to make at least the second best of both worlds, of agricultural productivity and wildlife conservation.

It is also encouraging to find that some modern farming methods do less damage than those which they are replacing. Ploughing, the traditional means of cultivation, is a drastic process, with profound effects on soil structure and chemistry and on all subterranean living organisms. It is done to produce a tilth for planting, and to control weeds. Work at Rothamsted Experimental Station in England over thirty years ago showed the surface-casting earthworms, which produced Charles Darwin's 'vegetable mould' (the stone-free well-textured surface layer in old pasture), were eliminated, and their populations might not recover for 70 years. Where cereals have been grown continuously for many years using only chemical manures, levels of organic matter have fallen and soil structure has suffered. Ploughing also uses a lot of expensive and scarce fuel.

Some farmers are now ceasing to plough, and are turning to direct drilling. A field of wheat is harvested, and the straw and stubble is burned. This burning, if improperly controlled, can do immense damage to the countryside, to hedges, woods and verges, and their wildlife, and bodies

Drill to till



KENNETH MELLANBY

like the National Farmers' Union are trying to discipline their members to accept a safety code. But burning has many agricultural advantages. It gives some control of diseases, weeds and pests, and returns a substantial amount of immediately-available nutrient to the soil. A burned field, with little trash on the surface, is easily cultivated, and rain does not produce a quagmire as it does after ploughing.

The field is usually left for some

weeks after burning, to allow weeds and shed grain to germinate. During this period it is a rich feeding ground for birds which obtain little food where stubble is ploughed straight after harvest. The surface is next sprayed with the herbicide paraquat, which kills the growth, but which is immediately immobilised. The sprouting grain is unharmed, and thus it is possible to drill the seed at the same time. This saves a lot of fuel. There has in the past been controversy about the effectiveness of direct drilling, as some farmers have obtained higher and some lower yields than with conventional ploughing. It seems now that good results are obtained by those who persist even after initial disappointment. After three or four years, even on heavy clay soils, the organic matter level increases (as several tonnes per hectare of roots rot down to produce humus) and the soil structure improves markedly. The soil fauna, including the delicate worms, also flourishes, and makes its contribution. The amount of herbicide used may be less than on ploughed fields, for some weeds are killed by the paraquat and the rapidly-germinating crop may smother others, including the usually troublesome couch grass.

Direct drilling is done to help the farmer to grow the most profitable crops, not to please the conservationist. It will not bring back our flower-rich meadows or the hedgerows once full of singing birds. But it may make the soil healthier and support a richer soil fauna. This is the basis of healthy agriculture and a healthy countryside. We, the conservationists, must now find how to encourage wildlife in this milieu.

news and views

Gene derepression in tumours

from Robert Shields

MANY non-endocrine tumours produce a wide variety of peptide hormones; perhaps the best known example is ACTH, (adrenocorticotropin) secreted by oat cell carcinoma of the lung (Rees *J. Endocr.* **67**, 143; 1975). The favoured explanation for this ectopic hormone production is that it results from 'gene derepression' which is claimed to be a frequent occurrence in neoplasia (Odell *New Engl. J. Med.* **297**, 609; 1977). This idea goes some way towards explaining many of the features of ectopic hormone synthesis. Quite often these hormones are larger than the normally occurring peptides (as happens with ACTH). These larger proteins are probably the prohormone precursor to the normal mature protein. Other ectopic hormones lack the modifications normally made before secretion; for instance the subunits of human chorionic gonadotropin (hCG) secreted ectopically by a variety of tumours lack the normal complement of carbohydrate. The gene derepression hypothesis could also explain why ectopic insulin production (which requires the tumour to execute the correct cleavage of proinsulin) is seen so rarely and why ectopic synthesis of non-peptide hormones such as steroids or thyroxine (which would involve the derepression of an entire enzyme pathway) has not been found.

But a number of observations are not fully explained by a simple gene derepression hypothesis. First, if gene derepression is common in neoplasia why do cells plump so often for hormones when there are 1×10^4 – 5×10^4 other genes to choose from? A partial answer may be that most of these genes code for housekeeping functions, over-production of these might be lethal. However, one might still expect luxury proteins to be derepressed with roughly equal frequency. Ectopic production of non-hormone proteins is in fact seen quite often, the best known examples are carcinoembryonic antigen (CEA) or α foetal globulin and

placental alkaline phosphatase. It is probable that many other proteins are synthesised ectopically, and that only the lack of clinical symptoms or suitable assays prevents their detection (Odell & Wolfson in *Cancer*, **81**, (ed. Becker F.) Plenum, New York, 1975).

In spite of this range of tumour products specific tumours tend to be associated with particular proteins: squamous carcinoma of the bronchus often produces parathyroid hormone (PTH) but seldom makes ACTH or MSH (melanocyte-stimulating hormone) whereas these hormones are relatively common in oat cell carcinoma (*Br. Med. J.* **1**, 1300; 1976). Why 'gene derepression' (if that is the explanation) occurs so non-randomly has not been satisfactorily explained.

The relative frequency of ectopic hormone production makes it worth asking whether ectopic protein production is unique to neoplasia and whether gene derepression is widespread in tumour cells. A recent report presents evidence for the presence of the β subunit of hCG in liver and colon from normal subjects (Yoshimoto *et al. Science* **197**, 575; 1977). Like ectopically synthesised hCG the protein lacked carbohydrate. Other experimenters have reported finding ACTH in non-involved lung tissue from patients with lung cancer (Bloomfield *et al. Clin. Endocr.* **6**, 95; 1977). These results suggest that ectopic hormone production may be widespread and occur in normal tissue as well. These findings also cast some doubt on the usefulness of ectopic hormones as tumour markers.

mRNA changes

As the techniques of nucleic acid hybridisation become more refined it is possible to ask more directly whether gene derepression occurs in tumour cells. If cDNA to total mRNA from transformed cells is cross-hybridised to normal cell mRNA any unhybridised cDNA should represent sequences present only in the transformed cell. If this remaining cDNA

is then back hybridised to transformed mRNA an estimate of the abundance and complexity of the new mRNA species can be obtained. When such experiments were done with human fibroblasts transformed by SV40 virus it was found that the great majority of mRNA species in transformed cells were present in the normal cells as well (Williams *et al. Cell* **11**, 901; 1977). At most around 3% of the mRNA in tumour cells was new (and part of this was probably virus coded) so vast host gene derepression did not occur. The estimate of how many new host cell sequences were derepressed can only be approximate but the authors calculate that around 35 new species may be involved. Since there are around 10^4 different mRNAs in the cell anyway 35 sequences does not represent large amounts of new information.

In these experiments care was taken to maintain normal and transformed cells under identical conditions to minimise the possibility that the differences seen were unconnected with transformation. Real tumours do not bother with these precautions and it might be expected that their mRNA patterns would diverge widely from normal. Surprisingly when mRNAs from HeLa cells was compared with those from normal cells there was very little difference once again suggesting that few new mRNA sequences were present in this abnormal cell. This result is particularly interesting as some clones of HeLa produce ectopic hCG (Lieblich *et al. Nature* **260**, 532; 1976; Ghosh *et al. Biochem. J.* **166**, 265; 1977). Whether hCG was produced in the clone examined by Williams is not known (Williams, personal communication), but in any event the number of new sequences in HeLa cells may be rather small.

These hybridisation experiments are only accurate in detecting entirely new mRNA transcripts. They are relatively insensitive to switching of mRNAs from low to moderate abundance classes. In other words, some mRNAs that are present in only a

few copies in normal cells may be present in tenfold greater amounts in transformed cells (and *vice versa*) and this would not be picked up. Viral transformation is known to affect the levels of mRNA for host proteins such as LETS and collagen (Adams *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 3399; 1977) and it would be surprising if other mRNAs remained unaltered.

Protein changes

Many experiments suggest a close relationship between the amount of mRNA and the amount of protein produced, so one indirect way to detect alterations in the levels of mRNA species is to examine the spectrum of proteins produced by normal and transformed cells by 2-D electrophoresis. There are relatively few differences in freshly transformed cells but when cells have been in culture for some time the differences are remarkable. Examination of the proteins from established lines of 3T3 cells, or 3T3 cells transformed with SV40 (DNA virus) or Kirsten sarcoma virus (RNA virus) revealed that although both transformed cell lines were similar to each other they differed from their untransformed counterparts in that around 30% of the proteins detectable in the transformed cells were not detected in untransformed cells (Strand & August *Proc. natn. Acad. Sci. U.S.A.* **74**, 2729; 1977).

Many of these 'new' proteins may not be new at all but represent increases in protein that were below

the level of detection in untransformed cells. Also some of these 'new' species may result from different patterns of glycosylation and other modifications in transformed cells which would alter the electrophoretic mobility of pre-existing proteins. Still other proteins may be induced because of differences in the utilisation of the culture medium by transformed and untransformed cells (Shill *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 3840; 1977). Nevertheless, the results show that these cells differ widely in the relative amounts of different protein produced and by inference in the relative amounts of many abundant mRNAs. If one can extrapolate from the experience with human cells (Williams *op. cit.*) (where few new abundant mRNAs were detected) it seems likely that the differing patterns of gene expression in normal and transformed cells may be due more to modulation of pre-existing mRNA sequences than introduction of new genes by oncogenic viruses or the derepression of pre-existing genes.

Whether this is generally true or not must await the preparation of cDNA probes to large numbers of different messages so that specific mRNA levels can be monitored directly after transformation and during evolution in cell culture. Perhaps a good way to start would be to prepare probes specific for ectopically produced proteins to see if mRNAs for these proteins are always present in normal cells (as the data of Yoshimoto *et al. op. cit.* might suggest) or whether these genes are really derepressed in tumour cells. □

then being made by comparing the rate of the drops' fall under gravity with the rise when an electric field is superimposed.

Millikan's original work involved studies of droplets of water or alcohol. On page 220 of his article in the *Philosophical Magazine* of February, 1910 he notes 'I have discarded one uncertain and unduplicated observation apparently upon a singly charged drop which gave a value of the charge . . . some 30% lower than the final value of e '.

This observation was remembered by many people recently when possible evidence for fractional charge was reported in the literature (La Rue *et al. op. cit.*). However the probable cause of Millikan's anomalous result is given by him only two paragraphs later in his paper: singly charged drops were so small that they evaporated rapidly during observation, hence a poor value of e was obtained. To avoid this evaporation problem Millikan ' . . . replaced the droplet of water or alcohol by one of oil, mercury or some other non-volatile substance . . . ' (*Phys. Rev.* **32**, 351; 1911). Never again did he see a fractional charge and one can safely conclude that evaporation was indeed the source of his spurious result.

If one intends to mount a search for quarks, whose occurrence as isolated particles sitting in lumps of matter is at best extremely rare, then one will want to study as much material as possible. The Millikan procedure is fine for very small drops where air friction can be used as a gauge for measuring the small forces on a charge in an electrical field. To study drops with masses of order 10^{-6} – 10^{-4} g an alternative technique is required. Hence studies turned to levitation of ferromagnetic or diamagnetic and superconducting materials by applying external fields. As an example of these experiments one can cite Gallinaro and Morpurgo (*Phys. Lett.* **23**, 609; 1966) who find no evidence for fractional charge in an experiment where more than 10^{-6} g of graphite were studied in a magnetic levitation spectrometer.

Possible evidence for fractional charge on a ferromagnetic material appeared in an experiment where twelve steel balls of 0.2 mm diameter were suspended in turn in a magnetic field. An electric field of $3,000 \text{ Vcm}^{-1}$ was then applied and the charge on the ball deduced. Of the 12 balls 10(!) had charges consistent with $\pm 1/3e$ and only one was consistent with integral charge. However, the authors (Garris & Zioc Nucl. Inst. Meth. **117**, 467; 1974) noted a source of systematic error, namely that due to

What price free quarks?

from F. E. Close

ALTHOUGH the concept of quarks as basic constituents of matter is almost universally accepted no evidence for isolated quarks has yet stood the test of time. However, interest in the search for 'free' quarks was heightened recently in the light of reports of new measurements of e , the magnitude of the element of electric charge, using variations of the Millikan oil drop technique (La Rue *et al. Phys. Rev. Lett.* **38**, 1011; 1977; Gallinari *et al. Phys. Rev. Lett.* **38**, 1255; 1977; Bland *et al. Phys. Rev. Lett.* **39**, 369; 1977). Quarks are widely believed to have fractional electric charge ($2/3$ or $-1/3$ of a proton's charge) and if they can exist as free particles then this property would make them easily

recognisable; hence the interest in the above experiments.

Measurements of e

In advance of Millikan, Ehrenhaft and others (refs in Millikan *Phys. Rev.* **32**, 349; 1911) had deduced the magnitude of the elementary charge from the average behaviour of charged particle swarms in electric and gravitational fields. Millikan appears to have been the first person to make a measurement by studying a single charged carrier, and the result quoted in his 1917 paper (*Phil. Mag.* **34** No. 199, July 1917) is within 1% of that accepted today. Every physics student has met the ideas behind this experiment whose essential feature is that extremely small drops of oil ($\approx 10^{-12}$ g) are observed in an electric field; an accurate determination of the charge

crystal structure and impurities metal surfaces are not equipotential surfaces. These 'patch effects' cause the potential of the metal surface to vary by 50 to 100 mV from place to place hence inducing a field gradient which can act on dipole moments induced in the balls by the applied electrical field. Reinserting several of the balls four months later the deduced charges had altered by as much as $0.45e$ showing that there is a long term change in the patch effect. The authors note that time-dependent changes in the patch effect could have caused a spread in the observed charge covering the measured range and so one cannot conclude that this experiment provides positive evidence for fractional charge.

Recent excitement was generated by reports of evidence for fractional charges on superconducting niobium (La Rue *et al. op. cit.*). Eight niobium balls each of mass $\approx 10^{-4}$ g were levitated in turn in a magnetic field. The charge on a ball can be changed at will by β^+ or β^- emitters adding positive or negative units of electronic charge to the ball. By studying the oscillations of a ball in the field the total force on the ball is deduced. The experimental problem which is crucial to interpreting the results is to decide the extent to which dipole forces might mimic a charge. Much of the paper is devoted to discussing these effects. The conclusion is that five niobium balls heated on a niobium substrate had charges compatible with zero whereas three balls heated on a tungsten substrate had charges of $(+0.337 \pm 0.009)e$, $(-0.001 \pm 0.025)e$, $(-0.331 \pm 0.070)e$ which, if genuine charge effects, are remarkably consistent with quark expectations.

If correct, it has been suggested that the earlier failures to find quarks might be understood if isolated quarks had some preference for existing on heavy nuclei. This does not seem to be borne out however as Bland and coworkers (*op. cit.*) have recently suspended tungsten particles in an electric field. A radioactive source can change the charge on a particle and the new electric field that balances the ball against gravity is determined. Sixty-nine balls were studied and no evidence for fractional charge on a ball was found. Over 10^{12} atoms of tungsten have been studied in this experiment and the authors plan to extend their study to other heavy elements.

Another null result has recently been reported by Gallinaro *et al. (op. cit. 1977)*. These authors studied iron balls of masses $\approx 10^{-4}$ g each and find no evidence for isolated quarks in over 10^{21} nucleons.

Ironically the observation of an isolated fractionally charged quark would

now cause more turmoil to theorists than its non-observation even though all are convinced that the quark concept is correct. Much effort is being devoted to see whether quarks can theoretically be confined inside hadrons in groups of three or as quark-antiquark pairs since these are what seem to be the rules in Nature.

A fifth quark, but no quarks

A recent discovery is that of a new particle, the upsilon (see *News and Views* 268, 297; 1977), which provides the first evidence for a possible fifth flavour of quark. The upsilon has

similar metastability to the ψ (*News and Views* 257, 535; 1975) but is approximately three times as massive (the ψ discovery in 1974 provided the first clear evidence for the existence of the fourth flavour—the charmed quark). It seems that the upsilon is a bound state of a new quark and its antiquark in analogy to the ψ being a bound state of a charmed quark and charmed antiquark.

Hence we have now evidence for five types (or flavours) of quark, yet the intensified search for an isolated quark reveals no sign of one. The quark paradox continues to deepen. $\square$

Rapid RNA sequencing

from Maria Szekeley

1977 has proved a fruitful year for nucleic acid sequencing. After the outstanding results with DNA achieved with new methods which shorten the time required for sequencing from years to weeks, these new principles have now been introduced to the temporarily neglected field of RNA sequence determination.

Several new developments have been reported at the same time. An RNA sequencing method based on the same principles as Maxam and Gilbert's DNA sequencing technique has been worked out simultaneously by Donis-Keller, Maxam and Gilbert (*Nucleic Acids Res.* 4, 2527; 1977) and by Brownlee's group in Cambridge (Simoncsits *et al.* (See this issue of *Nature*, page 833). This is also a rapid gel sequencing technique in which RNA fragments all ending at a given base are fractionated according to size so that the positions of this base in the RNA sequence can be read off the gel electrophoretic pattern. There is some difficulty involved in the production of these fragments from RNA. Splitting by base-specific endonucleases is used in place of the chemical reactions used for DNA. Although specific cleavages at purine residues are easy to achieve (RNase T_1 is specific for G, RNase U_2 for A in appropriate conditions), there are no enzymes available which split only at C or only at U residues. Simoncsits *et al.*, overcome this difficulty by using pancreatic RNase and RNase I from *Physarum polycephalum*. The former cleaves at C and U, the latter at A, G and U. By comparing the patterns obtained from the four enzymic digests it is therefore possible to deduce the positions of all four nucleotides, although with not quite 100% accuracy. The method may be improved; at its present stage the sequences obtained still need some

further confirmation. But it is rapid, sensitive and widely applicable and thus represents a very important new approach in RNA sequencing.

The other highly efficient gel sequencing technique—Sanger's plus and minus method (See *News and Views*, 267, 104; 1977) has also been adapted to RNA by Brownlee and Cartwright (*J. molec. Biol.* 114, 93; 1977) and has been successfully applied to three eukaryotic messengers: ovalbumin, globin and immunoglobulin mRNAs. The principle is the same in DNA and RNA sequencing, but different enzymes are used. cDNA is synthesised on the RNA template by reverse transcriptase and the same enzyme is used in the minus systems to extend these DNA fragments. Degradation of the DNA fragments in the plus systems is achieved by making use of a hitherto unknown property of DNA polymerase I: in the presence of Mn^{2+} ions this enzyme shows a 3'-exonuclease activity towards the DNA strand of a DNA:RNA hybrid.

There are still some difficulties: interference by RNase activity of reverse transcriptase, anomalies in copying regions of the template where long hairpin loops are present. Ambiguities could, however, be corrected and the above anomalies even proved useful in some cases to locate hairpin loops in the RNA. The accuracy of the technique is not always perfect, occasional artefacts may cause minor mistakes, as is the case also with Sanger's technique. The sequences obtained should therefore be confirmed by other sequencing methods. Still, the possibility of working out a 200-residue long sequence in one experiment, to at least 90 to 95% precision is a development which will greatly accelerate studies on RNA structure.

It is predictable that progress

will be fastest in the study of eukaryotic messengers, not only because these molecules are at the centre of interest, but also because Brownlee's method is especially suitable for messenger RNA sequencing. Reverse transcriptase requires a primer to synthesise cDNA. This is perhaps the only drawback of the method, as the synthesis of oligonucleotide primers is time-consuming and presupposes some knowledge of the RNA sequence. But the poly(A) tail of eukaryotic mRNAs facilitates this task: a 10-residue long oligo(dT) or more suitably an oligo(dT) to which two nucleotides, complementary to the 3' terminal dinucleotide of the mRNA, are attached, can anneal to the beginning of the poly(A) tail and enable the reverse transcriptase to start copying the mRNA exactly from the 3' terminus. Brownlee and Cartwright worked out in this way a 172-nucleotide long sequence from the 3' terminus of ovalbumin mRNA. Although they had to use 13 gels to find the best conditions and to correct the ambiguities, the sequence of the major part could be read off a single gel.

The most interesting recent result of the rapid gel sequencing techniques is the completion of the nucleotide sequence of a eukaryotic messenger: rabbit β -globin mRNA. The structure of this RNA species has attracted keen interest since the isolation of the first pure mRNA. The first attempt to determine the nucleotide sequence of globin mRNA was made years ago and since then practically all sequencing techniques have been applied to this RNA at one stage or another to elucidate some part of its structure. Progress was initially rather slow: the classical techniques had only limited application to large molecules which could not be labelled *in vivo*. Copying techniques were introduced; in dif-

ferent laboratories different enzymes were used to produce highly labelled complementary strands to mRNA which could be subjected to fingerprint analysis. In Cambridge, Proudfoot and Brownlee synthesised cDNA from globin mRNA with the aid of DNA polymerase I (*Nature* **252**, 356; 1974; *J. molec. Biol.* **107**, 491; 1976). At Harvard and Yale, Forget *et al.* (*Proc. natn. Acad. Sci. U.S.A.* **72**, 3614; 1975) and Marrotta *et al.*, (*Progr. Nucl. Acid Res. molec. Biol.* **19**, 165; 1976) applied a double transcription procedure, using reverse transcriptase on globin mRNA to obtain a cDNA transcript which in turn was copied into cRNA with *Escherichia coli* RNA polymerase. Salser's group at the University of California followed similar principles, sequencing parts of the cDNA and parts of the cRNA (*Progr. Nucl. Acid Res. molec. Biol.* **19**, 177; 1976). Lockard and RajBhandary at Massachusetts Institute of Technology used the 'wandering spot' procedure to determine the 5'-terminal capped structure (*Cell* **9**, 747; 1976). As a result of these different approaches major parts of the primary structures of rabbit globin and human globin mRNAs have been established.

In the April issue of *Cell* the long-awaited complete sequence of this mRNA has been published. Baralle (*Cell* **10**, 549; 1977) and Proudfoot (page 559) describe the 5' and 3' noncoding regions, respectively, applying the rapid gel sequencing technique of Brownlee and Cartwright. At the same time, Efstratiadis, Kafatos and Maniatis, using the gel sequencing technique of Maxam and Gilbert, determined the nucleotide sequence of their synthetic globin gene and thus also derived indirectly the sequence of globin mRNA (page 571). They present a 576-nucleotide long sequence almost complete in itself. The perfect agreement of this sequence with partial information available on the mRNA structure and

with the amino acid sequence of the protein proves that genes produced by *in vitro* transcription of mRNA and cloning of the double-stranded DNA product are faithful copies. □

Poland's 'fifth problem'

from I. J. Smalley

THE Polish government has defined nine major problems which are to be investigated by the country's scientists and engineers. These include such basic topics as developing a national electronics industry, finding a cure for cancer and tackling the difficult problem of water supply in a country with only one major river system. The fifth problem is housing—trying to find ways of overcoming the shortage of housing units in Poland. Such was the destruction caused in Poland during the Second World War that the housing balance has still not been restored.

Within project 5 perhaps the most urgent need is to find alternative sources of building materials, and this is closely followed by a need to improve the organisation of the building industry. The modern building material is concrete and for general purpose construction there is no real alternative. In a normal concrete only about 25% of the volume is occupied by the cement paste while the remaining 75% is occupied by inert space filling aggregate—and Poland is very short of aggregate.

Various approaches to the building material problem are being tried out at the Technical University in Warsaw. A. Degler is attempting to produce aggregates from pulverised fuel ash (PFA). Poland's coal-fired power stations (like those in the UK and USA) produce vast quantities of PFA and although some of this can be used as a pozzolanic cement substitute it would be best to convert it into aggregate. Degler's investigations on pelleting and extrusion processes have shown that it is possible—but there are delays in designing and constructing the necessary industrial capacity for the actual production. One of the nine major problems is that of environmental protection, so that if ways of using waste as aggregate can be found, two of the major problems will be alleviated.

Halina Badowska at the Technical University is attempting to find ways of utilising waste material from fertiliser

Maria Szekeley is a lecturer in the Department of Biochemistry at Imperial College, London.

THE rabbit globin mRNA sequence differs in various respects from the known nucleotide sequences of prokaryotic messengers. The latter contain a purine-rich tract near the initiation site, about 10 nucleotides towards the 5' end, which can form several base pairs with a pyrimidine-rich region at the 3' end of bacterial 16S RNA. No stable base-pairing is indicated between the 5' non-coding sequence of globin mRNA. A possible site for interaction of these RNAs was found around the initiation codon: here 6 base pairs could be formed. This, however, would imply that the initiator tRNA and the ribosomal RNA compete for the same

sequence in the mRNA.

The 53-nucleotide long non-coding sequence at the 5' end does not suggest formation of stable hairpin loops, but may allow base pairing with a stretch in the 3' noncoding region. A secondary structure which brings together the 5' and 3' ends of mRNA may be relevant to the control functions of the untranslated regions.

The 3' noncoding region is 95 nucleotides long. It shows extensive homology with the human β -globin mRNA. The human sequence has, however, an addition of 39 nucleotides which may have arisen by gene duplication.

I. J. Smalley is a lecturer in the Department of Civil Engineering at the University of Leeds.

production to make building blocks. Gypsum is the critical raw material and blocks using a form of gypsum cement and a polystyrene filler have been produced. A complex programme is also under way to develop resin materials as cements. There is in fact quite a severe shortage of Portland type cement in Poland although ironically enough Polish cement is exported. The need for foreign exchange is deemed greater than the need for cement.

G. Chrabczynski has been particularly concerned with the production of cement building units in which the hydration is accelerated by heating. This is often done under pressure in an autoclave but Chrabczynski has successfully developed methods which do not need the slow and expensive autoclave stage. Concrete railway sleepers with adequate strengths have been produced using much less cement than normal. Success, however, has been at the pilot plant level. There remains the difficult problem of organising factory production of the new items. Research is going on in Chrabczynski's section on organisation and management at all levels in the prefabricated building

block industry and it seems that, unless the persistent organisational problems are solved, the technical advances made in the laboratories will not benefit those who need project 5 to succeed.

It is ironic and regrettable that the new showpiece hotels in Warsaw have had to be built by Swedish contractors. These are certainly impressive buildings but it is a pity that desperately-needed foreign exchange had to be used to build hotels to give foreign businessmen and tourists the opportunity of returning the money to Poland. Project 5 is paradoxical in many ways. The population has recently passed 34 million—a level it reached for the first time just before the war, so that the colossal population losses have been replaced but the housing stock lags. Should family apartments have absolute priority or should the national construction industry be used to stimulate overseas trade? Major economic decisions will affect project 5: it is to be hoped that they will not have too adverse an effect. The 'fifth problem' does not have the scientific glamour of a possible cure for cancer but the Polish people would welcome its solution. □

New calculation of the nucleon optical potential

from P. E. Hodgson

THE differential cross section and polarisation for the elastic scattering of nucleons by nuclei can be very well described by a complex potential: the real part refracts the incident waves and the imaginary part absorbs them, just as a light wave is refracted and absorbed by a medium of complex refractive index. The absorbing potential takes account in a global way of all the reactions that remove flux from the elastic channel.

Unfortunately the potential describing the scattering is not unique: it is frequently possible to find several different potentials that fit the same data. This is unsatisfactory because these potentials give different wavefunctions, and unless we know which is correct it cannot be safely used to calculate the cross sections of non-elastic processes such as inelastic scattering and nucleon transfer reactions.

In principle the ambiguities between the potentials can be resolved by calculating the potential from more fundamental data, in particular from the nucleon-nucleon interaction,

together with the structure of the target nucleus. In practice, however, this is difficult due to the many-body nature of the problem.

Over the past few years there have been many attempts to calculate the optical model potential. These differ in the choice of approximations made to simplify the calculations and they give results in fair agreement with the experimental data. Particularly encouraging work has been carried out by Jeukenne, Lejeune and Mahaux of the University of Liège, who have been able to account for the volume integrals of the optical potentials (*Nature* **253**, 163; 1975).

Recently a new approach to this problem has been made by Brieva and Rook of the University of Oxford, and some preliminary results were reported at the Conference on Nuclear Structure at Tokyo in September. They start from the Hamada-Johnston potential that gives an accurate fit to the nucleon-nucleon scattering data, and then solve the Bethe-Goldstone equation for the motion of nucleons in infinite nuclear matter subject to this potential. This gives a quantity called an averaged *t*-matrix that is used to obtain the

optical potential by folding with the nuclear density distribution, using

$$V(r) = \int \rho(r') t(|r-r'|) dr'$$

with addition of exchange terms.

In this calculation the approximations were carefully chosen to facilitate the calculation without too much loss in accuracy. It differs from most previous work in giving both the real and imaginary parts of the optical potential, and allows the radically-dependent *t*-matrix to be non-local and energy-dependent. The spin-orbit potential is also calculated and is found to be very similar to the phenomenological spin-orbit potential.

The optical potential they calculated for ^{40}Ca is shown in Fig. 1 and compared with the phenomenological potential that gives the best fit to the differential scattering and polarisation data. It is particularly notable that the calculated real potential does not have the same radial shape as the phenomenological Saxon-Woods form. This indicates that this very frequently-used analytical form of the potential may not always be adequate. The calculated imaginary potential shows considerable structure, partly due to the form of the nuclear density distribution and partly due to the self-consistency requirement imposed on the potential. The differential cross section and polarisation calculated from these potentials are shown in Fig. 2. The agreement is strikingly good, indicating that it is now possible to calculate optical model potentials that give fits to the data that are comparable in quality to those found with phenomenological poten-

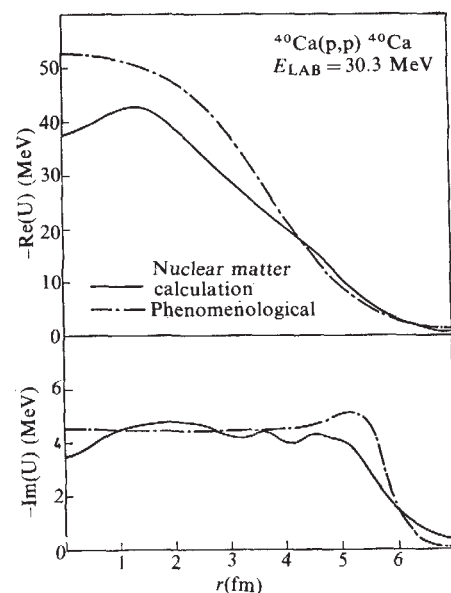


Fig. 1 Real and imaginary parts of the optical potential for protons on ^{40}Ca calculated from the nucleon-nucleon interaction compared with the corresponding phenomenological potentials.

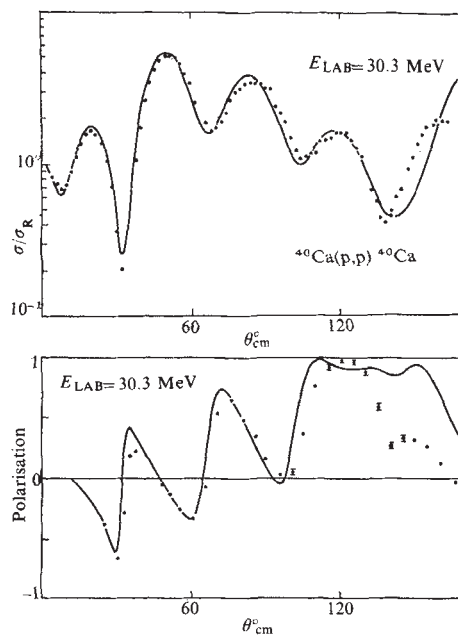


Fig. 2 Differential cross section and polarisation for the elastic scattering of 30.3 MeV protons by ^{40}Ca compared with optical model calculations using the calculated potential.

tials with many adjustable parameters. In some respects they are even superior as they can give features of the cross sections that cannot be explained by phenomenological potentials.

This work is very promising, and needs to be extended and applied to a range of nuclei to enable its validity to be explored. It should also prove possible to extend the method to the calculation of the optical potentials appropriate to composite particles such as deuterons and alpha particles, and perhaps also to heavy ions. The availability of optical potentials that can be reliably calculated from fundamental data instead of being obtained by phenomenological analyses of elastic scattering data will be of the greatest value in nuclear structure studies. □

Modulation of hormone receptors

from J. R. Tata

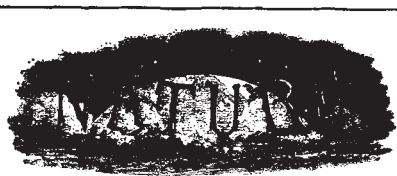
It is now widely accepted that receptors for hormones and other cellular signals, whether located on the cell's surface or intracellular, turn over rapidly so that the modulation of receptor level can be an important element regulating a cell's activity. About two years ago an account of the first reports on the possibility that some hormones themselves may control the levels of their own receptors or receptors for other hormones appeared in these columns (*News and Views* 257, 741; 1975). Whereas autoregulation of hormonal receptors is still a matter of some controversy, many investigators now accept that one hormone can regulate the level of receptor for another. In some instances it is hard to see the physiological relevance of such interplay among hormones but a recent paper suggests that it may be of clinical significance.

Endocrinologists and cardiologists have for long been aware of the tachycardia accompanying hyperthyroidism or thyrotoxicosis and that cardiac function in such patients was abnormally sensitive to small changes in adrenaline levels. Williams *et al.* (*J. biol. Chem.* 252, 2787; 1977) now explain this well-known syndrome by suggesting that thyroid hormones regulate the number of adrenergic receptor molecules in cardiac cells. In a carefully designed experiment they used the binding of the potent β -adrenergic antagonist (—)-dihydroalprenolol, to compare the number and affinity of β -adrenergic receptors in membranes isolated from myocardium of normal rats and animals made hyperthyroid by injections of the two thyroid hormones, L-thyroxine and triiodo-L-thyronine. Within 3–7 days, this treatment caused an increase in the binding capacity of cardiac membranes from 89 fmol of (—)- ^3H -dihydroalprenolol per mg protein in normal (euthyroid) rats to 196 fmol per mg protein. On the other hand, there was no difference in the affinity of the receptor for dihydroalprenolol or the β -adrenergic agonist isoproterenol (K_D at equilibrium of 2–15 nM dihydroalprenolol) between membrane preparations from euthyroid and hyperthyroid animals. Williams *et al.* therefore conclude that thyroid hormones control the number but not the nature of β -adrenergic receptors in myocardial cells. This would explain the tachycardia and increased myocardial contractility in clinical hyperthyroidism. Their observations would

also explain an earlier report that cultured foetal mouse hearts were rendered more sensitive by triiodothyronine to the chronotropic effect of β -adrenergic receptor stimulation (Widenthal *J. Pharmacol. exp. Therap.* 190, 272; 1974). If the same were true of other cells, then these findings could also explain skeletal muscle tremor or the enhanced rate of glycogenolysis and lipolysis in hyperthyroid subjects or experimental animals.

Are the observations of Williams *et al.* merely a peculiarity of thyroid hormones or are they part of a general pattern of an interhormonal regulatory device? Several recent publications point to the latter possibility. Such evidence comes mainly from studies on the interaction between hormones regulating reproductive processes which are known to be under multiple hormonal control. To cite but a few examples, Charreau *et al.* (*Molec. Cell Endocrinol.* 7, 1; 1977) have put forward a novel explanation why androgen, which like other steroids is not thought to act on adenylate cyclase, causes an increase in cyclic AMP levels in the rat prostate. These workers found that dihydrotestosterone (the active form of testosterone in target cells) increases the number of receptors for prolactin, a hormone known to influence male sexual development and to have a marked effect on adenylate cyclase in other tissues such as the mammary gland. Any interpretation of long-term studies of this kind in whole animals suffers from the drawback that one may not be measuring direct effects of hormones on receptor levels. For this reason, the demonstration that FSH (follicle stimulating hormone) induced receptors for hCG (human chorionic gonadotropin) directly *in vitro* in rat granulosa cells (Nimrod *et al. Nature* 267, 632; 1977) is significant. Hormonal interplay can of course operate in both directions, namely that one hormone can also dampen the sensitivity of a cell to another. Thus, it is interesting to note a recent report by Bhakoo and Katzenellenbogen (*Molec. Cell Endocrinol.* 8, 121; 1977) in which progesterone is shown to suppress the synthesis of oestrogen receptor molecules in the uterus. The resulting suppression of the replenishment of receptor molecules would thus explain the lowered uterine sensitivity to oestrogen that is known to be provoked by progesterone.

It is not difficult to predict that we shall be seeing many more examples of hormonal modulation of hormone receptors, particularly in reproductive tissues, mammary gland and the neuroendocrine complex, where multiple hormones are involved in the regulation of their development or adult



A hundred years ago

THE question has been discussed of late whether the ancient Greeks had an acute and true sense of colour. I remember once to have seen the remark that Sophocles shows his want of colour-sense by speaking of wine-coloured ivy. Now this really shows how true his perception of colour was. I enclose two ivy leaves which I have gathered to-day off a wall; I could have gathered plenty of the same colour, which, as you see, is claret colour.

JOSEPH JOHN MURPHY

Old Forge, Dunmurry,
co. Antrim, October 21
From *Nature* 16, 25 October, 551; 1877.

J. R. Tata is at the National Institute of Medical Research, London.

functions. The question then arises as to whether modulation of receptor levels is restricted to interhormonal regulation or whether this phenomenon is even more widespread. Raff (*News and Views* 259, 265; 1976) has reviewed the general topic of regulation of receptors, particularly where it concerns immunological sensitivity and lymphocyte maturation. A recent paper from Munck's group (Smith *et al.* *Nature* 267, 523; 1977) is therefore significant in this context since it suggests that non-hormonal signals could also modulate hormone receptors. These workers studied glucocorticoid receptors during mitogen (concanavalin A) stimulation of human lymphocytes and noted that blast transformation and mitosis is associated with a striking increase in the number but not nature of receptor sites per cell. How such a modulation of receptor numbers is brought about remains unknown, but their observations explain quite nicely the increased sensitivity to glucocorticoid hormones known to be brought about by mitogenic stimulation of lymphocytes and may also be relevant to the general question of how steroid hormones control the immune response. □

Pollen analysis at INQUA

from Paul Colinvaux

The X INQUA Congress was held in August, 1977 at the University of Birmingham. Pollen analysis was one of the many topics bearing on Quaternary research that were discussed.

FOR most of its existence, pollen analysis has been a tool of other disciplines; being used to age strata, to supply the temperatures of past times, to serve phytosociologists as they mapped the supposed movements of plant associations through millennia.

Time stratigraphy is now provided by Quaternary geology and isotope chemistry. Oceanic sediments now provide maps of past global climates. The press of contemporary ecological theory is to question strongly the validity of many former ideas about the integrity of plant communities. These developments allow pollen analysis to be used in more exhilarating ways, and the new ways are now taking hold.

One group of pollen analysts is interested in the ancestry of the remarkable complexity of the deciduous forests of Eastern North America. For several years this school has been

growing more confident in its insistence that the forest associations of Eastern America are comparatively recent assemblages of plants without the integrity, or persistence as communities, required by classical plant sociology. This school has used the pollen record to infer the responses of plant populations to known climatic change, borrowing climatic data from other disciplines. At the conference these analysts were confident that their case is made. Pollen analysis shows that plant associations beyond the front of the Wisconsin ice were different from all the associations we know today; and the tundras, prairies, and forests that replaced each other on the lands once held by ice were always critically different from the modern assemblages.

However, INQUA also saw reports of the refined use of pollen data to infer past climates. Fossil pollen assemblages are being compared with many surface samples, and the data are then being analysed with multivariate statistical techniques. It sounds like using principal components and canonical series to do what was once done by eyeball fits, but the approach is vitally different. The old way was to seek modern plant associations that were homologous with past communities and to graft on to those past communities the climate under which the supposed modern homologue was living. The new way is to select from ancient pollen spectra only those taxa found in comparable numerical abundance in surface samples, and then to compare the distributions of this select array with distributions of the same plants in historic times. There is in this no dubious assumption about the integrities or environmental requirements of whole plant communities; the only biological assumption is the respectable Darwinian one that niche requirements of each species do not change very much. Climatic reconstructions of this kind follow the methods that have been used by students of foraminifera in deep sea cores to reconstruct ancient water temperatures. INQUA saw several elegant papers using this approach. At INQUA, therefore, pollen analysis was used both to refute the existence of permanent plant communities and to improve the resolution of the climatic inferences that can be drawn from pollen.

All pollen analysts at INQUA drew comfort from the certainty that the oxygen isotope record of deep sea cores yields a chronology of the volume of global ice. The isotope record is not of ocean temperature, as was thought only a short while ago, but of glacial ice. It tells us nothing about world climate, other than the existence of ice itself, but it is a global time stratigraphy and

it must loosely relate to the particular climate of any one place. All pollen analysts have found themselves asking 'Are the local developments of vegetation that I see in step with the curve defining relative ice volume?' INQUA saw the first data relating isotope measurements to pollen in the same samples of offshore cores, revealing a gratifyingly close synchrony between changes in coastal forests and the volume of distant glaciers. A climatic linking mechanism is inescapable.

There were the expected developments in longer records and explorations in lesser-known parts of the world. Those with the long cores will look to the ice-volume curves to help with their time-stratigraphies and the first attempt of this kind to date long cores with the help of the ice volume curve was reported—a splendid long record from Searls Lake in California. The still very uncertain possibilities of remanent magnetism are also being explored to date pollen records back to the postulated Blake Event and beyond at Japanese Biwa, in Alaskan Imuruk Lake, and for the superb section from Grande Pille in France. The next INQUA may well see some ambitious reconstructions from these places describing the shufflings of plant species on single sites for more than 100,000 years.

The number of pollen records from the rest of the world grows slowly but steadily. Mapping the past in the Soviet Union continues as pollen analysts follow the completed map of the Quaternary deposits of that vast country, taking particular note of regions where archaeologists are active. But in lower latitudes where we need to know most there is still least to report. From the Indo-Pacific there are a few scattered records but from the South American equator near silence, particularly from the Amazonian basin. A member of the CLIMAP team at INQUA was heard to remark that he had had to infer the Wisconsin age vegetation of Amazonia from a paper on the modern distribution of bird species.

Yet in Africa there is progress, as papers come from Ethiopia and Chad, the lakes of the Rift, the dry basins of Sahara, and an array of stations in the South. But this African work shows how formidable are the difficulties of pollen analysis in tropical regions where the crucial species leave little pollen and where the species list is huge. As D. A. Livingstone remarked, deep in Tanganyika and the other lakes of the Rift lies a detailed history of the vegetation of this cradle of humanity since the Pliocene and beyond, but he despaired of seeing it read in his lifetime. □

review article

Claims and accomplishments of applied catastrophe theory

Raphael S. Zahler* & Hector J. Sussmann†

Several representative attempts to apply catastrophe theory to biological and social science problems turn out on close analysis to be characterised by incorrect reasoning, far-fetched assumptions, erroneous consequences, and exaggerated claims. Catastrophe theory seems to have made no significant contributions to biology and the social sciences, and to have no advantage over other better-established mathematical tools which have been used to better effect.

EMBRYOLOGY, ethology, ecology, and geology; physics, economics, dynamics, and linguistics; prison riots, literary symbolism, and the Vietnam war—these are some of the subjects to which catastrophe theory is said to be applicable. Its novel mathematical apparatus seems to be a near-universal tool, according to its proponents: "Properly understood and exploited, this ever-expanding web of concepts promises mankind a unique weapon against ignorance and a profound insight into the universe"¹.

We disagree. And because we feel that the many researchers now being attracted to catastrophe theory stand to gain nothing but disappointment and wasted time, we have written a critical study of applied catastrophe theory². Our conclusion is that the claims made for the theory are greatly exaggerated and that its accomplishments, at least in the biological and social sciences, are insignificant.

This is because catastrophe theorists have misused the basic mathematics in ways that lead to incorrect reasoning; they have offered models which are based on unreasonable assumptions and which lead to erroneous conclusions; and they have made predictions which are either vacuous, tautologous, vague or impossible to test experimentally. We do not say that catastrophe theory cannot possibly be applied. There may be legitimate uses in areas of physics and engineering, and some have suggested that the use of topological methods will be conceptually stimulating to researchers. So far, however, its record is poor.

We stress that we are not discussing the correctness or importance of catastrophe theory as a purely mathematical subject: we are sceptical only about its usefulness as a tool for extramathematical applications. We present this critical picture because we are excited about the prospect of new applications of mathematics, and concerned that many will be disenchanted with all of modern mathematics when they discover, as we have, that catastrophe theory is a blind alley.

This article is organised round a list of ten main kinds of defect found in catastrophe theory models in biology and social science. Each is illustrated briefly by one or two examples. For more detail, the reader should consult our longer paper². We are grateful for the advice of R. FitzHugh, J. C. Scanlon, H. Othmer, C. S. Hui, B. Goodwin, J. Sturtevant, G. Velicelebi,

S. Mabrey, J. F. Chlebowski, E. S. Crelin, J. F. G. Auchmuty, N. Van Arkel whose comments have been of great help in the preparation of this paper.

The cusp catastrophe

Most applied catastrophe theory is based on the 'cusp catastrophe' (Fig. 1). In this picture, the horizontal plane represents the possible values of two control parameters a and b , and the behaviour of the system is plotted on the vertical x -axis. For example, suppose we place a globular protein in solution and study its denaturation in terms of the concentration a of a denaturant and the temperature b . The extent of denaturation x_0 (as measured by the negative of the optical rotatory dispersion) as a function of denaturant concentration a_0 and temperature b_0 can be plotted as a point in the vertical line through (a_0, b_0) ; all such points form the curved 'cusp' surface shown, according to Kozak and Benham³.

So far this is just standard analytical geometry. The alert reader, however, may have noticed that above each point (a_0, b_0) in the shaded region R of the control plane lie not one but three different points of the behaviour surface; which one represents the actual behaviour at (a_0, b_0) ? The answer, given by the delay rule, is that it depends on how (a_0, b_0) is approached. As a and b vary continuously, x is to vary continuously in the cusp surface as long as possible. When continuous variation of x is not possible, then x is to jump to another sheet of the surface. For example, suppose the control parameters vary in such a way that the point (a, b) moves from Q_1 to Q_2 along the path P_1 , shown in Fig. 1. Initially, there is only one point in the cusp surface lying above each point in P_1 , so we follow the line shown from Q_1 towards J . Even when we pass point J the delay rule keeps us on the bottom sheet. But, when we reach point K , we must jump to K_2 on the upper sheet as shown. This is how sudden changes of behaviour in response to smooth changes of controls occur in catastrophe theory models.

Why the cusp? Catastrophe theorists argue that, according to a "deep mathematical theorem" of René Thom, essentially any system where sudden changes occur and where two control parameters appear can be described by the cusp; it is an inevitable, universal paradigm. According to Kozak and Benham, in fact, Fig. 1 correctly predicts the denaturation behaviour of RNase or collagen subject to the influence of temperature and CaCl_2 concentration. For example, following path P_1 from Q_2 to Q_1 , or P_3 from Q_3 to Q_4 gives a sigmoid curve for x , which corresponds to the appropriate experimental data. By distorting the cusp M somewhat (which is permissible

*104 Hubinger Street, New Haven, Connecticut 06511.

†Department of Mathematics, Rutgers College, Hill Centre for the Mathematical Sciences, Busch Campus, New Brunswick—New Jersey 08903.

in the theory), data from other systems can be modelled, they say.

Confusion about continuity

But when we look more closely at the model, things begin to come apart. First of all, are the transitions from one conformation to another really discontinuous jumps? The authors themselves admit that the experimental curves are not vertical, but this is only part of the problem. For a given small protein such as RNase, Van't Hoff's relation

$$\Delta H = RT^2 \frac{\partial \ln K}{\partial T}$$

implies that the enthalpy change ΔH sets a precise bound to how steep the temperature denaturation curve can be, even in a two-stage denaturation model (under the usual assumption that the molecules do not interact). Thus for a given protein

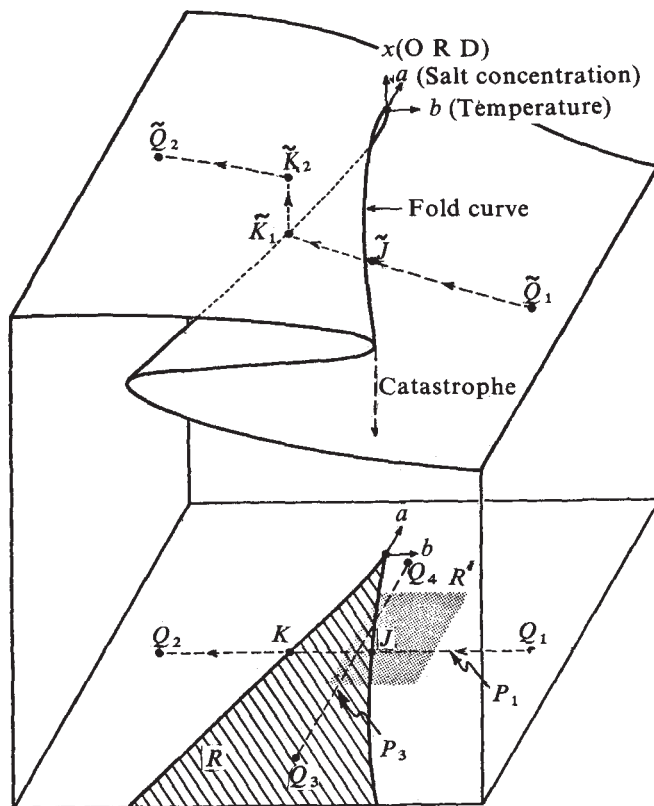


Fig. 1 An example of the cusp catastrophe, as it has been used to model protein denaturation.

we cannot even say that discontinuous jumps represent a limiting case. Now one of the main claims of catastrophe theory is that it is a unique way of applying mathematics to discontinuous phenomena. We have just seen that the reality of denaturation is inherently continuous; the same can be said for most biological situations that catastrophe theory has tried to model. It is therefore not surprising that most model-familiar to scientists are continuous. In these situations catastrophe theory gives no advantages in exchange for its greater complexity. Furthermore, in social science models, as we will see below, catastrophe theory often makes the reverse mistake, turning an obvious dichotomy into a continuous variable.

These continuity questions are not just mathematical fine points. By freely confusing the intuitive notion of 'jump' as a rapid change with the precise mathematical notion of a jump discontinuity, catastrophists are prone to construct false or

misleading arguments (see, for example, the discussion of embryology below).

Our next criticism is more serious. Having plotted experimental data that correspond to the part of the cusp above the region R' (Fig. 1), the authors claim that for the two-variable situation, "the mathematical theory states that if a change of morphology from one state to the other is observed", then the system must be described by a cusp. And even more: "if one has, say, two or three representative points from the bifurcation set of a given denaturation experiment, it should be possible to construct the surface M , and thereby to predict the behaviour of the system for values of the constraints not studied experimentally". Surely a theory that would allow a scientist to make two or three experimental observations on a laboratory system and then, with no other information or assumptions, deduce the behaviour of the system in all other conditions would be a great achievement. However, it is too good to be true: the authors have made serious blunders on at least three levels.

Use of Thom's theorem to justify extrapolation

The suggestion that Thom's theorem makes it possible to deduce the complete shape of a behaviour surface from some partial information about it is made by E. C. Zeeman⁴, the leading applied catastrophe theorist. (For example, "through catastrophe theory we can deduce the shape of the entire surface from the fact that the behaviour is bimodal for some control points".) However, Zeeman's statements are false. Take, for instance, the surface M of the denaturation model. Mathematically, it is simply not true that Thom's theorem forces M to look like Fig. 1. There might be many cusps or folds, and they need not be located at the origin nor oriented as shown. Furthermore, Thom's theorem cannot be used to infer the shape of an entire surface from some partial knowledge of it. Quite the contrary: any surface is arbitrarily close to a surface which satisfies the conclusion of Thom's theorem, and therefore (since all observations have a finite error) invoking this theorem can tell us nothing new about behaviour. This conclusion, which applies to all CT models, is not really surprising: suggesting, as Zeeman does, that one can deduce non-trivial facts about "the shape of all possible equilibrium surfaces" from a theory that makes no physical or behavioural assumptions is tantamount to claiming that the world can be deduced by pure thought—a claim which few scientists would accept.

Prediction contrary to fact

Biochemically, the extension of the surface from the region R' to the entire plane is not only mathematically unsupported but factually wrong, because it leads to implications that conflict with experimental data. For example, if we follow path P_1 in the right-to-left direction, we saw above that, according to the delay rule, this drop in temperature will cause renaturation at a temperature below that at which denaturation took place. But such hysteresis simply does not occur with collagen or RNase, the systems Fig. 1 is supposed to apply to. Thus the cusp is not a suitable model for the denaturation of these or most other simple proteins. It is true that there are some large proteins which exhibit hysteresis, but one cannot build a general theory of denaturation on atypical examples. In fact, the authors' attempt to extend their idea to other simple systems forces them to construct a model (Fig. 2) which is not even diffeomorphic to a cusp (they seem not to have noticed the vertical line above the origin), and therefore falls outside the realm of catastrophe theory altogether.

Lack of true testable predictions

Practically, the catastrophe theory analysis, even if it were correct, tells us nothing. Though the authors refer to the sigmoid curves obtained from paths P_1 and P_3 of Fig. 1 as 'predictions', this is quite untrue; they are merely graphs of the experimental data. Similarly, 'predictions' made by Zeeman in CT studies of embryology⁵ and neurophysiology⁶ are either (1) contained in

the data or (2) purely unverified hopes, or (3) independent of CT or (4) just wrong.

Consider, for example, Zeeman's studies of how homogeneous tissue of an embryo differentiates into two types separated by a frontier, a phenomenon he calls divergence or differentiation⁶. The 'main theorem' of this paper is worth quoting in full: "Homeostasis, continuity, differentiation, and repeatability imply the existence of a primary wave. In other words a frontier forms, moves and deepens, then slows up and stabilises, and finally deepens further".

It is hard to imagine what we could learn from such a breath-takingly vague statement, but we need not ponder this very long, since its proof is wrong as well.

Misuse of genericity

For one thing, the surface eventually chosen to model the cellular differentiation is not dictated by Thom's theorem but is obtained by a series of arbitrary choices which are justified—if at all—by appealing to 'genericity'. This mathematical concept is a way of excluding 'exceptional' or 'degenerate' cases. (For instance: 'generically, the axes of an ellipse are not equal' means that circles are an exceptional class of ellipse or, more intuitively, that a 'randomly chosen ellipse' is quite unlikely to be a circle.) In the course of his proof, Zeeman arbitrarily translates the vague word 'repeatability' into the precise concept of genericity. But then, all his theorem says is that if nothing exceptional happens, then the frontier moves.

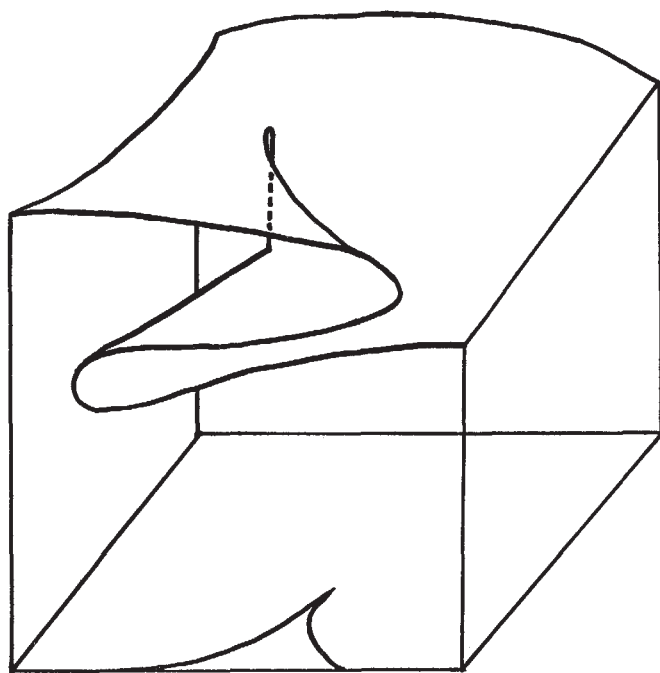


Fig. 2 Distorted cusp catastrophe; see ref. 3, Fig. 2.

Zeeman's 'proof' consists of no more than the observation that, if the frontier did not move, that would be quite exceptional. To evaluate this kind of reasoning properly, notice that the same logic, if correct, would apply equally well not just to frontiers of tissues but to anything whatsoever. So, Zeeman's reasoning 'proves' that everything moves except for those exceptional objects that do not. (For details, see ref. 2 §12 a).

Later in the same paper, Zeeman 'proves' that the frontier moves initially at constant speed. His proof is as follows: the frontier can be replaced initially, to first order, by its tangent, QED. Again, there is nothing in this reasoning that restricts its validity to frontiers of tissues, or to the time when the frontier forms. So Zeeman is really proving that everything moves

at constant velocity. But let us apply the same logic once again. Why not use the second degree Taylor approximation? And why not argue about the acceleration as Zeeman does about the velocity, and conclude that the acceleration is non-zero, because it would be exceptional if it were zero? This gives us the conclusion that everything moves with non-constant velocity.

Finally, why can we not take the zeroth order approximation? The conclusion then is that nothing moves at all.

We see that using these methods, everything, no matter how absurd, can be proved.

More than two thousand years ago, Zeno formulated his celebrated paradox of the arrow. Consider an arrow in its flight. 'At each instant of its flight the tip of the arrow occupies a definite position. At that instant the arrow cannot move, for an instant has no duration. Hence, at each instant the arrow is at rest. Since this is true at each instant, the moving arrow is always at rest. This paradox is almost startling. It seems to defy logic itself.'

The development of Calculus has resolved this paradox, and taught us how it is possible for things to change and, at the same time, be what they are at each instant. The position of the tip of the arrow is what it is at each time, but it is different at different times. Hence the arrow can have a definite position at each time, and not be at rest. What is true of the position is also true of velocity. The velocity has a value at each point, but this does not mean that it is constant. Zeeman, a Twentieth Century Zeno, is reformulating the paradox as a corollary, thus ignoring more than two millennia of mathematics.

But Zeeman's main theorem is wrong for other reasons as well: it is easy to show by means of counter examples that Zeeman's conclusions that the frontier does not stabilise or deepen do not follow from his assumptions. When these counter-examples were brought to the attention of catastrophe theorists, the theorem was defended by disclosing new interpretations of the terminology and the hypotheses. For instance, the hypothesis of 'differentiation' is now taken to mean that after some time two different types of tissue are found and no more changes occur. If so, then after some time the frontier no longer moves, so the frontier does indeed stabilise. But of course this 'proof' of stabilisation is transparently circular.

In general, the hypotheses on which many catastrophe-theory arguments are based are often vague. When terms like 'differentiation' and 'repeatability' are used without any precise translation, the correctness of the 'proofs' is hard to assess. Catastrophe theory papers consistently violate one of the most basic rules of the scientific method: state clearly what you mean and do not change definitions in the middle of your reasoning.

Misleading use of mathematics

The most basic criticism of Zeeman's catastrophe theory of embryology, however, is the way the mathematical foundations are used. Others have pointed out that the 'deduction' that differentiation will occur as a pattern of primary and secondary waves arises not from catastrophe theory itself but from the conditions assumed to exist before the primary differentiation process begins (for example a gradient of cell states). In fact, then, most of the results of this paper, regardless of their merit, have nothing to do with catastrophe theory at all.

To see why this is important, we must remind the reader of what catastrophe theory is and is not. First, events in which apparently 'catastrophic' changes occur are not necessarily 'catastrophes' in the technical mathematical sense. Catastrophe theorists agree that the term 'catastrophe' is reserved for certain kinds of singularity of smooth maps, seven of which have been described and classified elegantly by Thom. The keystone of catastrophe theory, in fact, is Thom's profound theorem. It is what is supposed to give catastrophe theory its deductive powers, because it allows one to conclude that a given situation must be represented by a cusp, or one of the other elementary catastrophes.

As we have illustrated, however, most catastrophe theory models make no use of Thom's theorem. Thus the cusp is no longer inevitable or unique.

How, then, does the cusp arise? Often, as in the embryology models, the cusp arose because the hypotheses were carefully chosen to make sure that it would. In other cases, however, experimental data are plotted and found to resemble a cusp^{3,6}. This is not surprising: if we plot an ordinary hysteresis loop in

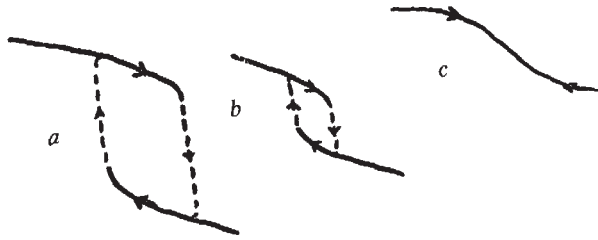


Fig. 3 Hysteresis loops of decreasing size.

two dimensions (Fig. 3a), in which fast vertical jumps alternate with slower horizontal movements, and then imagine a continuum of such loops getting smaller (Fig. 3b) and finally vanishing to form a continuous curve (Fig. 3c), we get a surface resembling the cusp of Fig. 1. Thus we can regard the cusp as a nice way of picturing a system with variable degrees of hysteresis. But it is certainly not unique in this: a surface with a double fold whose projection was any smooth curve (a parabola, for example: see Fig. 4) would do as well. So we must ask again: what is special about a model that looks like a cusp? Although some have found such pictures aesthetically satisfying, or a source of insight, we can only state that the fact that certain data happen to form a cusp-like shape tells us nothing new about the system.

Careless discussion of evidence

Catastrophe theorists have asserted that there is experimental evidence for some of their models. For instance, Zeeman⁴ writes, "I have constructed CT models of the heart beat, the nerve impulses and the formation of gastrula and of somites in the embryo. Recent experiments by J. Cooke and T. Elsdale appear to confirm some of my predictions". The facts are as follows. (1) No experiments have been made to test the predictions of Zeeman's nerve impulse models. As for the heart-beat, Zeeman is said to have made experiments in 1972, but the results have not been published (Zeeman, personal communication). (2) The catastrophe theory nerve impulse models disagree with experimental voltage-clamp data⁹ in several important aspects, deny the universally accepted concepts of the sodium leak and the independence of sodium and potassium channels, and lead to the wrong propagation speed for the action potential. (3) Zeeman's embryology paper⁶, besides being mathematically wrong, betrays the author's inexperience in embryology. For example (p. 27), Zeeman likens the embryonic neural tube to a roll of stiff paper which tries to maintain its curl. But experiment shows that cut neural tube persistently tries to unroll¹⁰. (4) T. Elsdale *et al.*¹¹ write: "... we do not yet conclude that the observations here presented have confirmed Cooke and Zeeman's model to the exclusion of others". (5) J. Cooke (personal communication) writes: "I, at least, do not regard any of the predictions of the model in which I am involved as being deeply distinctive to catastrophe theory".

Stewart¹ repeats the untrue assertion that Zeeman's embryology predictions have been "recently verified by experiment". Regarding the use of catastrophe theory in social sciences, Stewart writes: "Although most such models still lack precise data, an interesting exception is a study by Zeeman and several collaborators of how tension and alienation among prison inmates influence disorder. A cusp catastrophe fits the data

very well, the sudden jumps being riots and truces". In fact, when Zeeman and his coworkers plot points in a plane, and look for a cusp curve which fits them, they do not succeed in fitting one curve, so they use two, and claim that the cusp must have moved during the process. Finally, they do not use any statistical techniques to determine whether or not their pair of cusp curves gives a good fit (compare Fig. 5).

Dixon Jones⁸ has made a catastrophe theory model for budworm infestation. His model relies on a 'cusp' which is really not a cusp in the correct mathematical sense. His model predicts a fast rise of a variable, but the data rise slowly. Jones then rescues his model by making a logarithmic change of scale!

At this point the reader may suspect that we have chosen the weakest catastrophe theory papers as targets for our criticism. The contrary is true: the best biological applications are contained in the papers just cited. Others^{12,13} consist of long expositions of the mathematics of catastrophe theory, copious explanations of elementary biology, and vague speculations on how the two could be brought together.

Unreasonable or ambiguous hypotheses

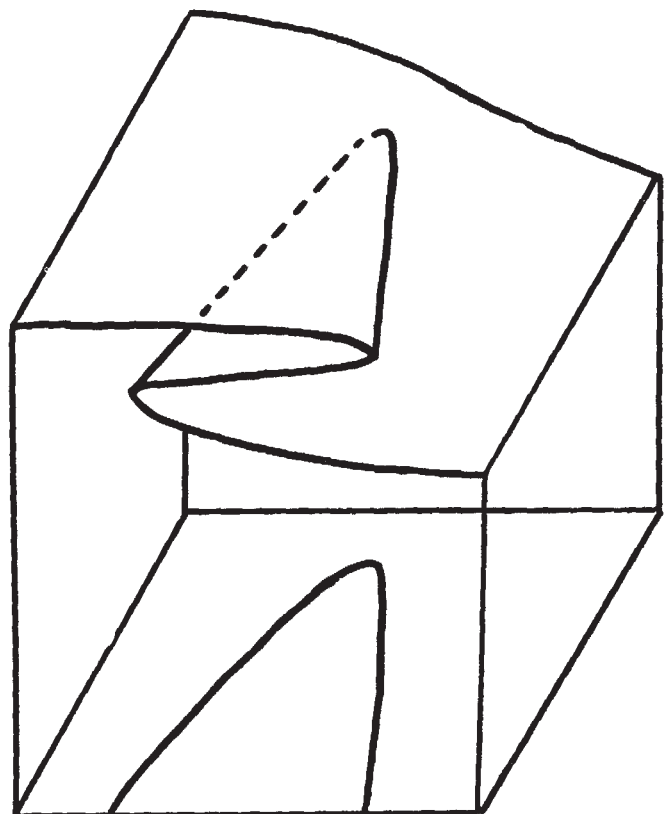
Isnard and Zeeman¹⁴ have a model for how a country makes decisions about going to war. It is assumed that public opinion depends on the 'cost' of the war, and the (perceived) 'threat' posed by the enemy nation.

Among the five hypotheses of the war model are: Hypothesis 2: If the cost of the war is low, then opinion will be unified, and the greater the threat, the greater will be the level of military action called for.

Hypothesis 3: If the cost is high, and the threat moderate, then opinion will be split between doves and hawks.

Now in a given conflict situation, no sensible person would advocate a level of military action so low that the enemy is sure to win an easy victory. Either one wants no war at all, or

Fig. 4 A folded surface which is not a cusp since the projection of the fold curve on the control plane does not have a sharp point.



one wants to use as much military force as he thinks is needed to win. (War will follow when at least one side miscalculates, and both sides think they can win.) So, the dichotomy war-no war is a discrete one, and this is so whether the cost is high or low. But, if this is taken into account, we are left with a better theory which, however, does not involve a cusp.

Not only are the hypotheses of the war model unreasonable but they do not lead to the conclusions reached by the authors. Their main claim—that the hypotheses imply, via Thom's theorem, that the behaviour set is a cusp surface—is once again dubious, since the proof, which is not given, is presumably

ingless. The same reasoning applies to the 'level of military action' in the war model, as we have shown.

Better alternatives

Defenders of catastrophe theory claim that it provides, for the first time in history, a mathematical method for modelling discontinuous phenomena. This is false because, as we have shown, catastrophe theory does not lead to satisfactory models. But it is also false because it ignores the whole body of discrete mathematics, the study of shock waves, bifurcation theory, and the mathematics of quantum mechanics. Better alternatives certainly exist.

Besides, no mathematical theory, no matter how elegant it may be, can serve as a substitute for the hard work of learning the facts about the world. Catastrophe theory is one of many attempts that have been made to deduce the world by thought alone. It offers to mathematicians "the hope of applying mathematics without having to know anything but mathematics"¹⁶. An appealing dream for mathematicians, but a dream that cannot come true.

Appeal of the theory

Why has catastrophe theory acquired such widespread popularity? One possible reason may be its impressive claims of universality and usefulness, backed up by a large number of (mostly unrefereed) publications praising each other extravagantly; another might be the unusual nature of the mathematics used, combining concepts that are completely inaccessible to anyone who is not a professional mathematician with the use of some pictures of amazing simplicity, so that the result is simple to grasp intuitively but difficult to criticise.

It may be said that catastrophe theory is a new theory and that, while all the applications proposed so far have serious flaws, each one has benefits that make it worthwhile. Our criteria for judging a theory (or method) are rather generous. Even if a theory rests on questionable assumptions, or is based on faulty reasoning, or leads to false conclusions, or deals with ambiguous concepts, or does not make true testable predictions, we are ready to accept that it may be valuable, so long as it does not have all these faults at once.

The catastrophe theory models that we have examined, however, combine all these faults. The assumptions on which they are based are unreasonable and/or vague. The reasoning used to draw conclusions from the hypotheses is mathematically incorrect. The conclusions are either trivial (for example "a frightened dog, if angered, may attack"¹⁴) or false.

The possibility that, in the future, catastrophe theory may produce solid applications, cannot be dismissed *a priori*. However, its spectacular failures should suffice to raise serious doubts. For a method that has been said to have "the potential for describing the evolution of forms in all aspects of nature", existing evidence is indeed disappointing. The scientific community must remain sceptical until the proponents of catastrophe theory succeed in substantiating their claims. The burden of the proof is on them.

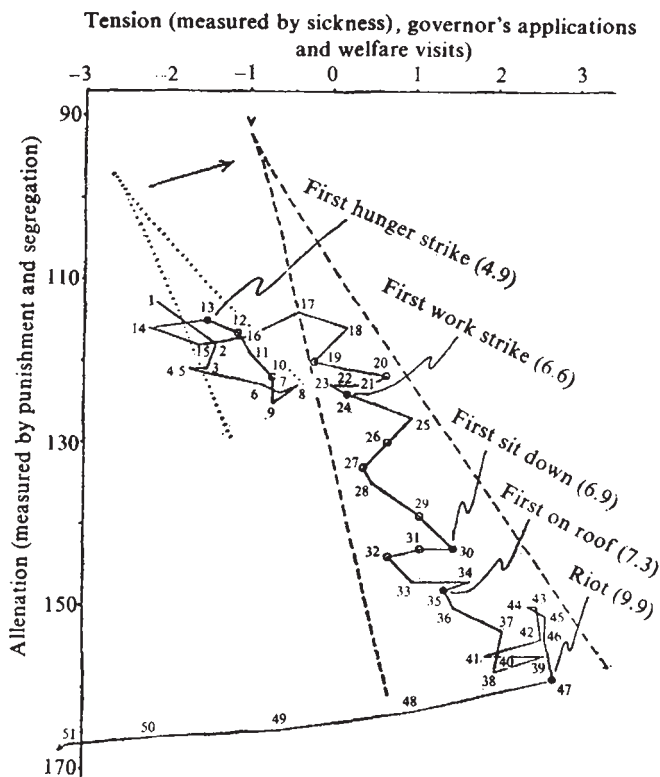


Fig. 5 The 'analysis of the data' of the paper by Zeeman *et al.* on prison riots¹⁷. These data are supposed to show that a cusp catastrophe gives a good fit.

based on some translation (also not given) of the vague hypotheses into precise language. It is not hard to show, however, that depending on one's interpretation of the hypotheses, either the conclusion does not follow; or the result does not depend on Thom's theorem, and is essentially contained in the hypotheses.

In fact, the only possible part Thom's theorem might play is to exclude a rounded bifurcation curve (such as the parabola of Fig. 4) in favour of a cusp (Fig. 1). But since either Fig. 1 or Fig. 4 would lead to the conclusions stated by Isnard and Zeeman, there is really no need to use Thom's theorem after all.

Spurious quantification

Catastrophe theorists often attempt to make a discrete variable into a continuous one so that CT can be applied. In Zeeman's dog aggression model⁴, for example, the level of aggression of a dog is considered as a continuous variable x , ranging "from outright retreat through cowering, avoidance, neutrality, and growling and snarling to attacking". He also states, however, that an 'attack catastrophe' occurs when the value of x jumps upwards from one sheet of the behaviour surface to another. Despite this ambiguity, it can be shown that with either interpretation the attack is embedded in a continuous family of behaviours. But this is absurd: the idea of a 'semi-attack' by a dog, or of a snake gradually attacking a person, is utterly mean-

1. Stewart, I. N. in *Encyclopedia Britannica Book of the Year* (Encyclopedia Britannica, New York, 1977).
2. Sussman, H. J. & Zahler, R. S. *Synthese* (in the press).
3. Kovak, J. J. & Benham, C. J. *Proc. natn. Acad. Sci., U.S.A.* 71, 1977 (1974).
4. Zeeman, E. C. *Sci. Amer.*, 234, 65 (1976).
5. Zeeman, E. C., in *Lectures on Mathematics in the Life Sciences* 7, 69 (1974).
6. Zeeman, E. C. in *Towards A Theoretical Biology* 4, 8 (Edinburgh University Press, Edinburgh, 1972).
7. Kline, M. *Mathematics in Western Culture* (Oxford University Press, New York, 1974).
8. Jones, D. D. *The Application of Catastrophe Theory to Ecological Systems* (International Institute for Applied Systems Analysis, Schloss Laxenburg, Austria, 1975).
9. Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* 117, 500 (1952).
10. Crelin, E. S. *J. exp. Zool.* 120, 547 (1952).
11. Eldsdale, T., Pearson, M. & Whitehead, M., *J. Embryol. exp. Morph.* 35, 625 (1976).
12. Woodcock, A. E. R. *Ann. N.Y. Acad. Sci.* 231, 60 (1974).
13. Dodson, M. M. *Math. Biosci.* 28, 243 (1976).
14. Isnard, C. A., Zeeman, E. C., in *Use of Models in The Social Sciences* (Tavistock, London, 1974).
15. Blainey, G. *The Causes of War* (Macmillan, New York, 1973).
16. Kolata, G. *Science*, 196, 287 (1977).
17. Zeeman, E. C., *et al. Br. J. math. stat. Psych.* 29, 73 (1976).

articles

Radio structure of 3C147 determined by multi-element very long baseline interferometry

P. N. Wilkinson & A. C. S. Readhead

Owens Valley Radio Observatory, California Institute of Technology, Pasadena, California

G. H. Purcell*

National Radio Astronomy Observatory, Charlottesville, Virginia

B. Anderson

University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, UK

We have determined the radio structure of the quasar 3C147 from multi-baseline VLBI data at 609 MHz using both a conventional method and a technique which uses the 'closure' phase information to produce a good approximation to a synthesis map of the source. The structure is similar to the central part of M87, with a bright core and a linear 'jet' of projected length ~ 1.5 kpc which is concentrated in bright 'knots'.

THE quasar 3C147 ($z = 0.545$, $V \approx 17$ m) is one of the more intense radio sources in the sky at decimetre wavelengths. Previous interferometer and interplanetary scintillation observations¹⁻⁶ have shown that its radio emission is confined to a region $\lesssim 1.0''$ in extent. At high frequencies most of the radiation comes from a complex region of the source $\lesssim 0.2''$ in size, while at low frequencies ($\lesssim 100$ MHz) the size of the predominant emission region is $\sim 0.7''$. The interferometer observations, in which only fringe amplitudes on single baselines were measured, were inadequate to determine in any detail the morphology of the compact region. We report here the results of more extensive observations which do reveal this structure in considerable detail. They constitute the most reliable complex radio structure determined by VLBI techniques.

The observations

The data were obtained in two independent observing runs, both at 609 MHz, using circularly polarised primary feeds on each antenna. In each case the data were recorded with NRAO MkII VLB equipment and the video tapes correlated on the NRAO processor⁷. The details of the telescopes used and their rf performance are given in Table 1. The six interferometer baselines are shown schematically in Fig. 1a, and the resolution achieved by each interferometer pair as a function of hour angle is shown in Fig. 1b.

The first observations, made in December 1973, involved only the three US stations (giving the three inner tracks in Fig. 1b) and were analysed solely by G. H. Purcell in 1974. The plots of fringe amplitude as a function of interferometer hour angle are shown in Fig. 2a. Further observations of 3C147 were made, using all four telescopes, in March 1975. We observed the source for nearly all the time for which it was jointly visible at each station in order to obtain the maximum possible 'closure' phase information, as well as visibility amplitudes.

A closure phase, C , is formed by the addition of the observed interferometer phases around a closed loop of baselines at a given time. The simplest loop is a triangle. Four triangles can be formed from our four station network (see Fig. 1a), but only three of these yield independent information. On each baseline we measured the fringe phase, ϕ_i , given by

$$\phi_i = \psi_i + \theta_i$$

where ψ_i is the phase of the visibility function and θ_i is a

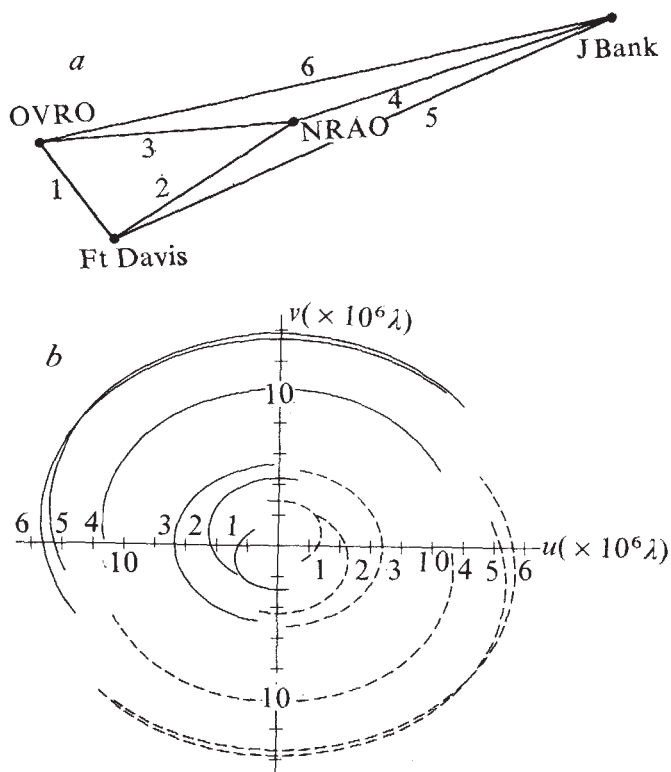


Fig. 1 a, Schematic representation of the six interferometer baselines. We formed closure phases around all four triangles but used mainly C_{123} , C_{245} and C_{346} (see text). b, The baseline of each interferometer as a function of hour angle. The source was observed continuously along each track. The baselines are as follows: 1, Ft Davis–OVRO; 2, NRAO–Ft Davis; 3, NRAO–OVRO; 4, J. Bank–NRAO; 5, J. Bank–Ft Davis; 6, J. Bank–OVRO.

*Present address: Jet Propulsion Laboratory, Pasadena, California

perturbation due to changes in the path delay to each telescope through the atmosphere and ionosphere, to phase drifts in the independent local oscillators and to uncertainties in the baseline and source position. A combination of these effects, with the ionosphere having a dominant role at this relatively low frequency, precludes the direct use of the six interferometer phases. It has been shown^{8,9} that in the closure phase the perturbations cancel, and thus the three closure phases we form

$$C_{123} = \phi_1 + \phi_2 - \phi_3$$

$$C_{245} = \phi_2 + \phi_4 - \phi_5$$

$$C_{346} = \phi_3 + \phi_4 - \phi_6$$

are equivalent to

$$C_{123} = \psi_1 + \psi_2 - \psi_3$$

$$C_{245} = \psi_2 + \psi_4 - \psi_5$$

$$C_{346} = \psi_3 + \psi_4 - \psi_6$$

respectively. The three closure phase curves we obtained are shown together with the six fringe amplitude curves in Fig. 3a, b and c. Unfortunately, because of a malfunction in the NRAO processor, the signal-to-noise on all baselines except NRAO-Ft Davis was reduced by a factor of two from the normal value.

The agreement between the independently measured fringe amplitude curves for the US baselines at the two epochs is within the expected uncertainty due to receiver noise and calibration errors (~5%). This agreement shows that there have been no significant changes in the structure of the source during an interval of 15 months¹⁰.

The brightness distribution

The brightness distribution of the source was inferred by Purcell from the 1973 amplitude data alone, using a model fitting method in which the parameters of elliptical Gaussian components are adjusted by an iterative procedure to produce a best fit to the data in the least squares sense. The complexity of this source is such that 13 gaussian components were required to produce the solid lines in Fig. 2a, which are the fringe amplitudes calculated from the final model. We do not list the parameters of this model, but rather display the contours of radio brightness which they represent; this 'picture' of 3C147 is shown in Fig. 2b. Note that because no phase information was included there is a 180° ambiguity in its orientation.

To analyse the more extensive 1975 data we devised a new approach which is an extension of the following scheme proposed by M. H. Cohen: beginning with a model of the source based on amplitude data alone, calculate the visibility phase of this model on all baselines and use this phase, with the observed amplitudes, to make a 'map'. Using the CLEAN technique¹¹⁻¹³ reduce the sidelobe level on this 'map' and thereby produce a new model of the source. The phases predicted by this new model are different from the original phases and thus if we repeat the process using these new phases the 'map' will be different from the original one. Iterate, using the new model to predict the phases each time and see whether the method converges to a stable 'map'. This method is essentially that proposed by Fort and Yee¹⁴, except that where we use the CLEAN technique, they set all negative regions of the 'map' to zero at each iteration. We decided to extend the procedure proposed by M. H. Cohen by including the closure phase data, and thus to make use of all the information available in present-day VLBI observations (that is fringe amplitude, closure phase and extent of (u,v) coverage).

Table 1 Interferometer elements

Location	Institution	Diameter (m)	System Noise (K)	Sensitivity (kJy ⁻¹)
Jodrell Bank, Cheshire, UK	Manchester Univ.	76	140	0.95
Green Bank, W. Va., USA	NRAO	43	160	0.21
Ft. Davis, Texas, USA	Harvard Univ.	26	300	0.08
Big Pine, Calif., USA	Caltech	40	260	0.18

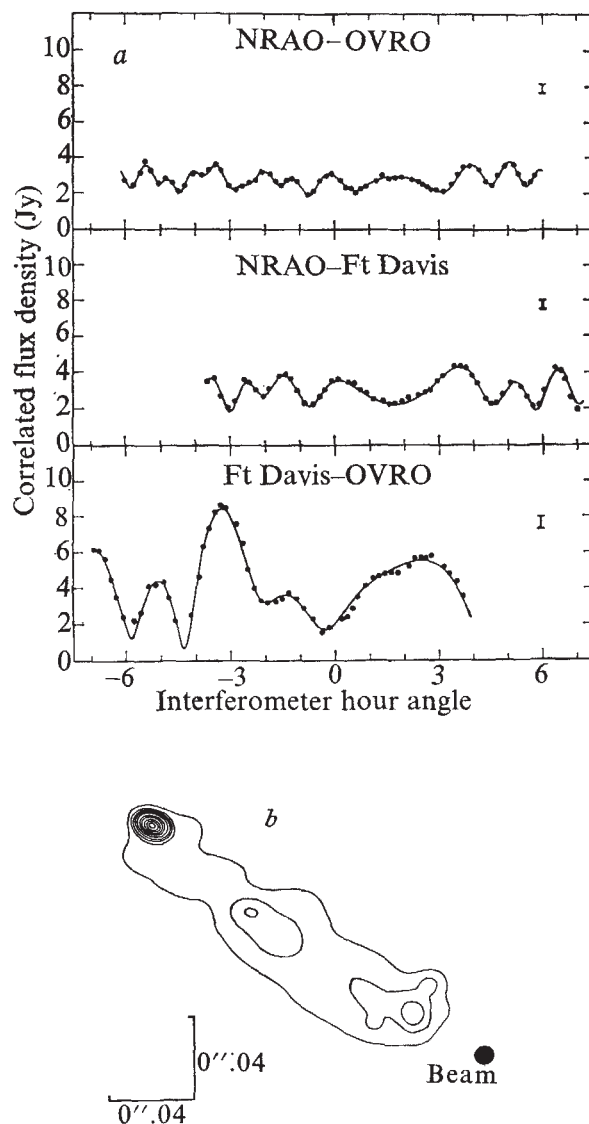


Fig. 2a, Fringe amplitude data on the US baselines in December 1973. Mean error bars shown include a contribution from the uncertainty of the flux scale on the different baselines. The solid lines are the predicted amplitudes from the model in Fig. 2b. b, Contours of radio brightness derived from the 13-component model fitted to the fringe amplitudes in Fig. 2a. The contour levels were chosen to correspond with those in Fig. 4. The original Gaussian components have also been convolved with a circular Gaussian (FWHM = 0.009") to simulate the effect of the 'clean beam' in Fig. 4.

This combination of the closure phase with an iterative 'cleaning' technique has subsequently proved to be a powerful method for determining source structure provided the (u,v) plane is sufficiently well sampled. The method will be described in more detail elsewhere, but the analysis of the 3C147 data provides a simple and illuminating example of its use. It is clear from the fringe amplitude and closure phase plots that this source contains a compact region which is barely resolved on the three longer baselines. It is reasonable to assume that the phase of the visibility function is almost constant and close to zero on these baselines (that is, $\psi_4 \approx \psi_5 \approx \psi_6 \approx 0^\circ$) irrespective of the detailed structure of this region. For each uv interval we can therefore solve the three independent closure phase relationships and determine unambiguously, within the accuracy of our assumption, the true visibility phase on the shorter baselines (that is, ψ_1, ψ_2 , and ψ_3). In this case we are using the compact region as a phase reference.

Note that the phase closure information has enabled us to

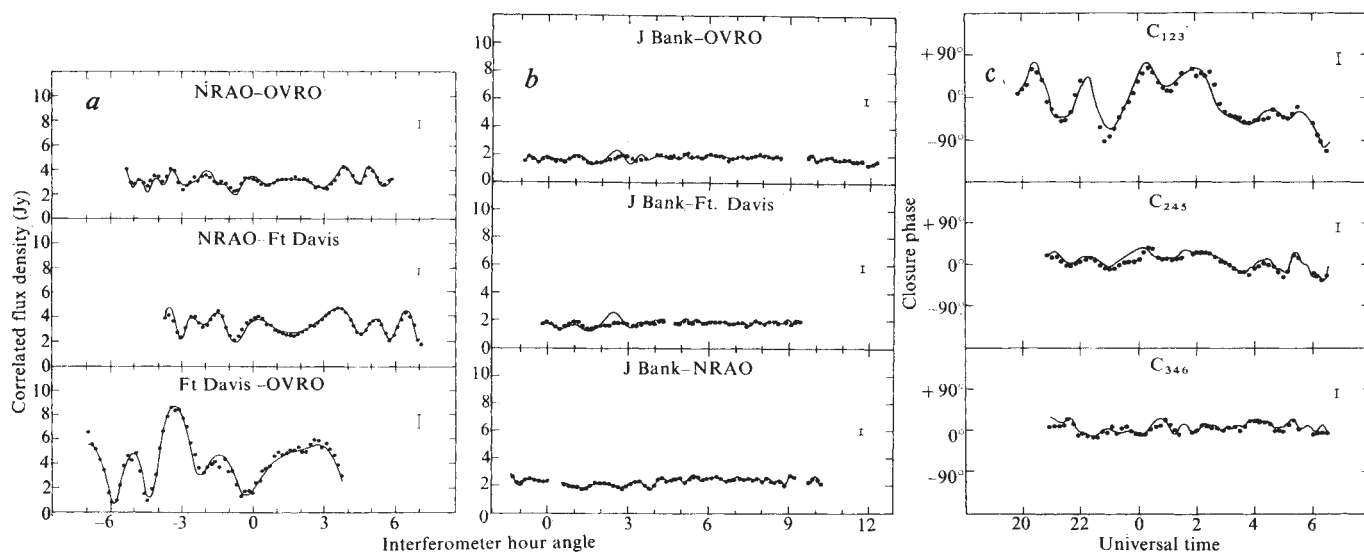


Fig. 3 *a, b*, Fringe amplitude data from March 1975. The signal-to-noise ratio on all baselines, except NRAO-Ft Davis, is only one half the expected value (see text). The solid lines are the predicted amplitudes from the brightness distribution shown in Fig. 4 *a, c*. Three independent closure phase relations from the March 1975 observations. The solid lines are predicted phases from the brightness distribution shown in Fig. 4 *a*.

resolve the usual 180° ambiguity in the position angle of the source, but it does not allow us to locate the absolute position of the radio source on the sky⁸. Apart from the absolute position of the source, we have derived all the complex visibility information accessible to these interferometers. In other words, the data can be regarded as synthesis observations of 3C147 with (u,v) coverage limited to the three US baselines.

The measured fringe amplitudes and the constructed phases were convolved on to a regular grid in (u,v) space using a Gaussian convolution function. Zero values were assigned to grid points not filled by this procedure. This array was then transformed into the sky plane resulting in the so-called 'dirty map' of the source, which is badly disturbed by sidelobes resulting from the incomplete u,v coverage. The effect of these sidelobes was reduced by means of a version of the CLEAN technique developed by D. Rogstad. The primary output of this program is an array of point sources in the sky plane which approximates the true source brightness distribution. This first 'map' can be improved slightly because of the small extent to which the extended structure does influence the observed fringe amplitude on baselines 4, 5, and 6. The corresponding phases can therefore be expected to show minor deviations from the assumed value of zero. To take account of this, the major features of the first 'map' were used to re-predict the phases on the long baselines and the analysis was repeated to produce a second 'map'. This process was continued and, while no significant changes were noted after the second or third iteration, five iterations were performed to check the stability of the derived brightness distribution.

The array of point sources produced by CLEAN has been smoothed with a Gaussian 'clean beam' whose FWHM is the same as that of the central maximum of the dirty beam. Fig. 4*a* shows the final composite clean map, the result of the fifth iteration, in which the clean beam is circular and has a FWHM of $0.009''$. In Fig. 3*a-c* the solid lines are the amplitudes and closure phases obtained by Fourier inversion of the final array of point sources only. The agreement with the data indicates that we have subtracted most of the real features of the dirty map.

We have also checked the sensitivity of this process to our original supposition of zero phases on the longer baselines. To do this, we repeated the above procedure exactly, except that initially we used phases predicted by the model of the source

shown in Fig. 2*b*. The result is shown in Fig. 4*b*; the brightness distributions in Fig. 4 are very similar and this gives us confidence that they accurately represent the compact features in the source.

The similarity between the brightness distributions shown in Figs 2*b* and 4, obtained from independent data by entirely different methods, is remarkable. The distribution of Fig. 2*b* predicts amplitudes lower and more smoothly varying than those actually observed on baselines 4, 5 and 6, indicating that the true distribution contains more compact features than those of the Gaussian model. Nevertheless, the agreement supports the conclusion of Pauliny-Toth *et al.*¹⁵ that the main features of a complex brightness distribution can usually be reconstructed reliably even without phase data. The objective nature of our method, however, the speed with which a complicated brightness distribution can be determined, and the avoidance of the 180° position angle ambiguity make this approach a significant advance on conventional model-fitting methods.

As well as the compact structure discussed above, 3C147 also has an extended component which has been observed both at low frequencies, by interplanetary scintillation observations^{8,16}, and at 2,695 MHz by single baseline, conventional interferometric observations⁴. At low frequencies ($\lesssim 100$ MHz) the extended component dominates, and $\gtrsim 80\%$ of the radiation must come from a region $\sim 0.7''$ in extent. At the higher frequency the compact features dominate, and only $\sim 25\%$ of the radiation originates in the extended region. In November 1976, 3C147 was observed twice with the Cambridge 5 km telescope at 15 GHz (M. Ryle, personal communication), and these observations show that the source is 30% resolved in p.a. $30^\circ \pm 5^\circ$ on the longest baselines, indicating a size $\sim 0.7''$. The source is unresolved in p.a. $\sim 120^\circ$. The position angle and the size are consistent with the model-fitting results of Donaldson and Smith⁴, and in Fig. 5 we show schematically a composite of our VLBI results with their preferred interpretation. The exact details of the extended structure are not known, but it is clear that the components are not colinear, and it seems that the outer lobe does not form a smooth extension of the jet. Note however, that only one contour of a Gaussian is drawn to indicate this outer lobe and it would not be surprising if future observations, with baselines of a few hundred thousand to about one million wavelengths, do

Table 2

Component	S_{peak} (Jy)	ν_{peak} (MHz)	θ_{eq} (arc s)	θ_{obs} (arc s)	Total energy (erg)	Pressure (dyn cm ⁻²)
Outer lobe	70 ± 10	120 ± 30	0''.35	~ 0''.6 × 0''.15	5 × 10 ⁵⁷	~ 10 ⁻⁸
Jet	20 ± 8	420 ± 100	0''.05	~ 0''.2 × 0''.025	5 × 10 ⁵⁶	~ 3 × 10 ⁻⁷
Core	3.2 ± 0.5	500 ± 200?	0''.012	~ 0''.006		
*Core (if on straight part of spectrum)	3.2	609	~ 0''.0055	~ 0''.006		

* See text.

in fact show a bridge of emission between the jet and outer lobe. In this case the source would resemble 3C66B (ref. 17). The minimum resolution of our observations ($\sim 2 \text{ M}\lambda$) is too great to study the outer lobes, although they are probably the cause of the systematic deviations of the expected fringe amplitudes from the data on the shortest (OVRO-Ft Davis) baseline. The flux density of these lobes must be relatively greater at 609 MHz than at 2,695 MHz because the compact features shown in Figs 2b and 4 only account for about half the total flux density of 35 Jy from the source at 609 MHz.

The radio spectrum of 3C147

The total spectrum of 3C147 has a maximum at $140 \pm 20 \text{ MHz}$, and a marked low frequency cutoff, which is thought to be due

to synchrotron self absorption¹⁸. We have constructed three-point spectra for the individual components using the total spectrum and high resolution observations at 609 MHz (present data), 1,660 MHz (refs 5, 19), and 2,695 MHz. The 2,695 MHz points are taken from a re-analysis⁵ of the data of Donaldson and Smith⁴. In all cases, except 609 MHz, the core was taken to be unresolved ($\lesssim 0.01''$).

The spectra of the individual components cannot be determined with great accuracy; however, it is clear from the total spectrum that the spectra of both the outer lobe and the jet must have low frequency cut-offs. We have fitted to the data a series of three-component spectra, which are consistent with the total spectrum (Fig. 5). The parameters of the different components derived from these spectra are shown in the first two columns of Table 2.

The spectrum of the core is the least well determined of the three components. There is no indication of a low frequency cutoff and it is not possible to infer, even approximately, the frequency at which this occurs. The value of ν_{peak} in Table 2 of $500 \pm 200 \text{ MHz}$ has therefore been included merely as an indication of what might be reasonable (see later).

The energy balance in the different components

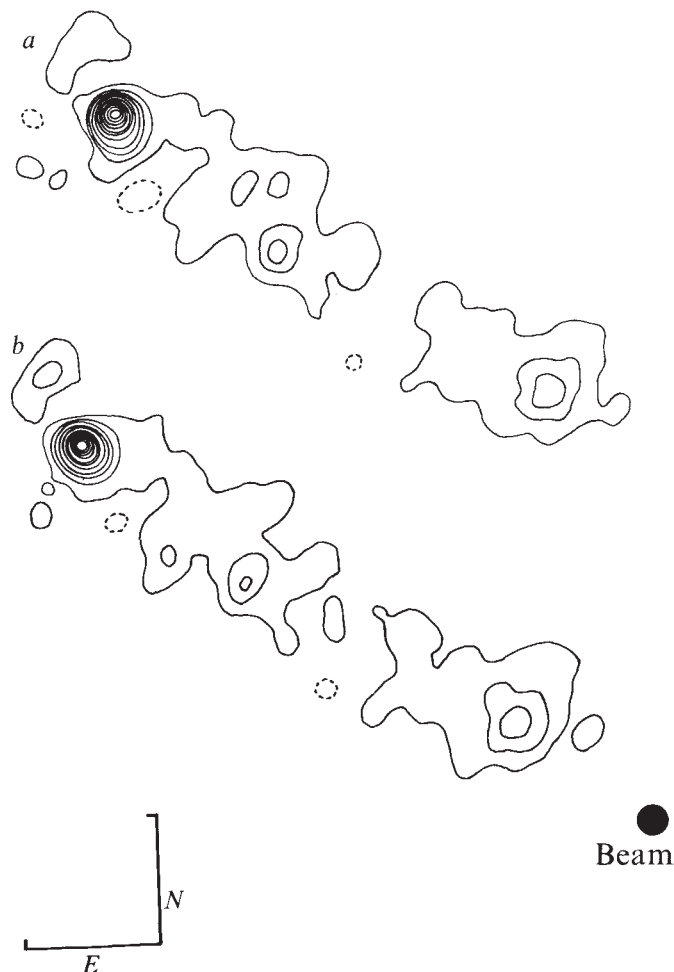
We now discuss the possible physical conditions which obtain in the different components.

The jet and outer lobe. A number of different mechanisms could be responsible for the low frequency cut-offs which we observe in the jet and outer lobes of 3C147. Of these, synchrotron self-absorption seems to be by far the most likely explanation since both free-free absorption and the Tsytovich-Razin effect would require too much thermal plasma in the components (~ 2 and $\sim 3,000 \text{ electrons cm}^{-3}$ respectively). If there were such a high density of thermal plasma present, it would cause considerably more depolarisation within the source than is observed²⁰. Scott and Readhead¹⁸ have recently shown that the turnover in the total spectrum is almost certainly due to synchrotron self-absorption (their work is based on low frequency measurements and therefore refers to the outer lobe), because the measured diameter at 81.5 MHz is within a factor 1.5 of the equipartition diameter, θ_{eq} , assuming that synchrotron self-absorption is responsible for the low frequency cut-off. Following the treatment of Scott and Readhead for a simple slab source we have calculated θ_{eq} for the three components of 3C147, and these are given in Table 2. It is clear that the observed angular sizes of both jet and outer lobe are consistent with equipartition.

The total energies and the pressures in the outer lobe and jet, calculated assuming equipartition between the particle and magnetic field energy densities are also shown in Table 2.

The compact core. The spectrum of the core is apparently flat (spectral index ≥ 0.2) down to $\sim 600 \text{ MHz}$. Since the cut-off has not been observed, we calculate a firm lower limit to θ_{eq} by assuming that the $S \propto \nu^{5/2}$ line of the self-absorbed spectrum passes through the 609 MHz point (see Table 2). The observations are only just consistent with equipartition ($\theta_{\text{eq}} \gtrsim 0.0055''$) if the core is a single homogeneous source. If the cut-off in the spectrum occurs at a frequency substantially lower than 609 MHz then there must be considerably more energy in

Fig. 4 Brightness distributions determined from the March 1975 amplitude and closure phase data using the method described in text. The 'clean beam' is a circular Gaussian with FWHM = $0''.009$. Contours are plotted at $-25, 25, 75, 125, 175, 275, 375, 475, 575, 675, 775, 875 \times 10^8 \text{ K}$. ($25 \times 10^8 \text{ K} \equiv 62 \text{ mJy}$ per unit beam area). The arms of the L each represent $0.040''$ arc. *a*, Initial phases on baselines 4, 5, and 6 predicted from a point source. *b*, Initial phases predicted from the model shown in Fig. 2b.



the particles than in the magnetic field. A flat spectrum can also be produced by several individually self-absorbed components whose spectra peak at different frequencies. The characteristic undulations can easily be missed if the individual flux density measurements are not determined to sufficiently high accuracy. Preuss *et al.*²¹ have recently detected at least two weak (maximum correlated flux density ≈ 0.26 Jy) components in 3C147 at 5,000 MHz with a baseline of $\sim 10^8$ wavelengths. Their

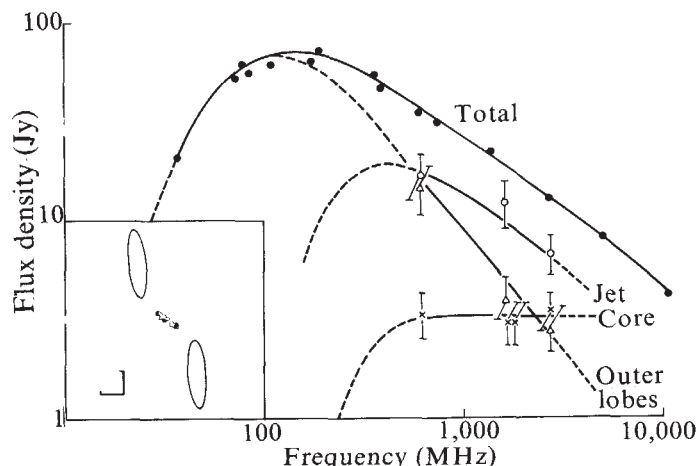


Fig. 5 Total source and individual component spectra. Errors of $\pm 20\%$ have been assigned to the component flux densities to take account of model fitting uncertainties. (See text for references.) (Inset: composite brightness distribution showing schematically the relationship between our results and the preferred interpretation of Donaldson and Smith⁴ at 2,695 MHz. The contour level in the outer components is arbitrary and the contour is drawn at the FWHM of their Gaussian model. Scales represent $0.15''$.)

much greater resolution and lack of knowledge of the location of these components makes it difficult to compare these results with ours. The characteristic size they derive, however ($0.004''$, that is, ≈ 30 pc) suggests that the results refer to the core and that our measured size may not refer to a single component.

Discussion

The observations and analysis described above constitute the most reliable determination of radio structure on the scale $0.01''$ to $0.25''$ for any source. The 'core', 'jet' and 'outer lobes' morphology is very similar to 3C66 (ref. 17), M87 (ref. 22) and 3C219 (ref. 23), and suggests that in trying to explain 3C147 one should concentrate on theories involving the generation of energy in a central highly active region, and the transfer of the energy to outlying regions via a collimated beam of waves and/or particles²⁴.

The angular extent of the jet ($\sim 0.20''$) corresponds to a projected physical length of ~ 1.5 kpc at the distance of 3C147. This is similar to the projected length of the M87 jet. Another similarity is that the 3C147 jet also contains discrete 'knots'. The two knots in 3C147 have linear diameters ~ 150 pc, somewhat larger than the radio knots in M87 (~ 50 pc) (ref. 25). Turland²³ has shown that there is a wide range in both the luminosity and size of radio jets. Our results on 3C147 are remarkable if the same physical processes are responsible for producing all jets, since the radio luminosity of the 3C147 jet is four orders of magnitude greater than that of the M87 jet, and yet the physical sizes of both jet and knots are very similar in the two cases. The difficulties associated with the confinement of the knots in M87 have been discussed by Okoye²⁶ and by Turland²³, who suggested that the optical extent of the knots is determined by the lifetime (~ 20 yr) of the electrons responsible for the optical emission assuming it to be synchrotron radiation.

The possibility, however, that these knots are confined should be seriously considered for the following reasons. Wilkinson²⁵ has shown that the sizes of the radio and optical knots in M87 are very similar, yet the radio emitting electrons have a lifetime $\gtrsim 10^8$ years, much greater than the light travel time across a knot. Similarly in 3C147, the synchrotron lifetime of the radio-emitting electrons ($\gtrsim 3 \times 10^8$ yr) is much greater than the light travel time across a knot, and therefore they must be confined. The pressure in the 3C147 knots is, however, at least an order of magnitude greater than in the optical knots of M87, and therefore it should be easier to confine the M87 knots than those in 3C147.

Our knowledge of the outer lobes is very limited but we do know that the overall position angle is different from the jet. Further VLBI observations with resolutions in the range $0.1''$ – $2.0''$ could determine the structure on this scale if the closure phase were used. The uncertainty about the exact locations of the lobes illustrates one limitation of the 'closure phase' technique—namely that we cannot use it to determine absolute positions. If we knew the position of the core it would be a simple matter to subtract the core and jet from the 15-GHz map of the source made with the 5-km telescope, and this would then show clearly the disposition of the more extended features.

We have shown that it is possible to use VLBI observations to make a reliable determination of the structure of compact features in a fairly complex source. To determine that structure on other scales we have also made three station observations at 327 MHz and 5 station observations at 1,667 MHz (in preparation). These should enable us to determine both the structure on a scale $\sim 0.002''$ to $\sim 0.5''$ and the spectra of the different components, since the inclusion of the closure phase information allows us to study the components at different frequencies without ambiguity. Thus these observations should provide us, for the first time, with a complete picture of the structure of a high redshift radio source on a scale of parsecs to kiloparsecs.

We thank Professor M. H. Cohen for useful discussions and support. We also thank Dr D. Rogstad for his help in using the CLEAN procedure, and H. Yee for computing assistance. We are grateful to the staff of each observatory for their help in making these observations, and we thank Professor Sir Martin Ryle for allowing us the use of 5-km telescope data before publication. A.C.S.R. gratefully acknowledges the support of the Royal Society Weir Research Fellowship. The Owens Valley Radio Observatory is supported by NSF grant AST 73-74660. The National Radio Astronomy Observatory is operated by Associated Universities Inc., under contract with NSF.

Received 17 June; accepted 9 August 1977.

1. Anderson, B. A. & Donaldson, W. *Mon. Not. R. astr. Soc.* **17**, 81–94 (1967).
2. Clark, B. G., Kellermann, K. I., Bare, C. C., Cohen, M. H. & Jauncey, D. L. *Astrophys. J.* **153**, 705–714 (1968).
3. Clark, R. W., Broten, N. W., Legg, T. H., Locke, J. L. & Yen, J. L. *Mon. Not. R. astr. Soc.* **146**, 381–397 (1969).
4. Donaldson, W. & Smith, H. *Mon. Not. R. astr. Soc.* **151**, 253–258 (1971).
5. Fort, D. N. thesis, Univ. Manchester (1970).
6. Readhead, A. C. S. & Hewish, A. *Memoirs R. astr. Soc.* **78**, 1–49 (1974).
7. Clark, B. G. *Proc. I.E.E.E.* **61**, 1242–1248 (1973).
8. Jennison, R. C. *Mon. Not. R. astr. Soc.* **118**, 276–284 (1958).
9. Rogers, A. E. E. *et al. Astrophys. J.* **193**, 293–301 (1974).
10. Readhead, A. C. S., Wilkinson, P. N. & Purcell, G. H. *Astrophys. J.* **215**, L13–L16 (1977).
11. Hogbom, J. A. *Astron. Astrophys. Suppl.* **15**, 417–426 (1974).
12. Schwarz, U. *Astron. Astrophys.* (in the press).
13. Schwarz, U. *Astron. Astrophys.* (in the press).
14. Fort, D. N. & Yee, H. K. C. *Astron. Astrophys.* **50**, 19–22 (1976).
15. Pauliny-Toth, I. I. K. *et al. Nature* **259**, 17–20 (1976).
16. Duffett-Smith, P. J. & Readhead, A. C. S. *Mon. Not. R. astr. Soc.* **174**, 7–17 (1976).
17. Northover, K. J. E. *Mon. Not. R. astr. Soc.* **165**, 369–379 (1973).
18. Scott, M. A. & Readhead, A. C. S. *Mon. Not. R. astr. Soc.* **180**, 539–550 (1977).
19. Kellermann, K. I. *et al. Astrophys. J.* **169**, 1–24 (1971).
20. Conway, R. G. *et al. Mon. Not. R. astr. Soc.* **168**, 137–162 (1974).
21. Preuss, E., Pauliny-Toth, I. I. K., Kellermann, K. I. & Shaffer, D. B. *Astron. Astrophys.* **54**, 297–300 (1977).
22. Turland, B. D. *Mon. Not. R. astr. Soc.* **170**, 281–296 (1975).
23. Turland, B. D. *Mon. Not. R. astr. Soc.* **172**, 181–189 (1975).
24. Blandford, R. D. & Rees, M. H. *Mon. Not. R. astr. Soc.* **169**, 395–415 (1974).
25. Wilkinson, P. N. *Nature* **252**, 661–663 (1974).
26. Okoye, S. E. *Mon. Not. R. astr. Soc.* **165**, 393–402 (1973).

Secular changes in marine inundation of USSR and North America through the Phanerozoic

A. Hallam

Department of Geological Sciences, University of Birmingham, P.O. Box 363, Birmingham, UK

Analysis of palaeogeographic and facies maps of the USSR and North America confirms an overall trend towards marine regression through the Phanerozoic with a major deviation in the late Mesozoic. These are related to world-wide changes of sea level attributable primarily to a combination of continental thickening controlled by subduction processes, and changes in the overall length and hence volume of the oceanic ridge system.

AN apparent secular withdrawal of the sea from the continents during the course of the Phanerozoic, superimposed on shorter-term oscillations of sea level, has been inferred from world palaeogeographic maps^{1,2} and held to support the notion of slow earth expansion of $\sim 0.5 \text{ mm yr}^{-1}$ during that time³⁻⁶. This interpretation has been challenged by Armstrong⁷ and Hallam⁸, who argue that the secular withdrawal can be adequately explained without invoking a change in the Earth's radius, and by Wise⁹, who disputes the very fact of secular change. The problem is re-examined here using the analysis of palaeogeographic and facies maps of the USSR and North America, which leads to a discussion on the relative merits of the various hypotheses that can be put forward to account for Phanerozoic sea-level changes. Attention is devoted to the longer-term secular changes rather than the shorter-term transgressions and regressions.

Soviet Union

The series of palaeogeographic and facies maps of the Soviet Union published in the 1960s¹⁰, as yet unmatched for any other extensive region, allow a detailed analysis of changes in the areal extent of sea through the whole Phanerozoic by a technique outlined elsewhere¹¹. All palaeogeographic maps involve, of course, a degree of inference, but there can be little doubt from the abundant data provided that the Russian maps give a close approximation to the area of continent flooded at several intervals for each successive geological period, although areas of present continental shelf cannot be included because of lack of data.

As shown in Fig. 1, the sea spread slowly in the early part of the Palaeozoic to cover approximately 60% of the

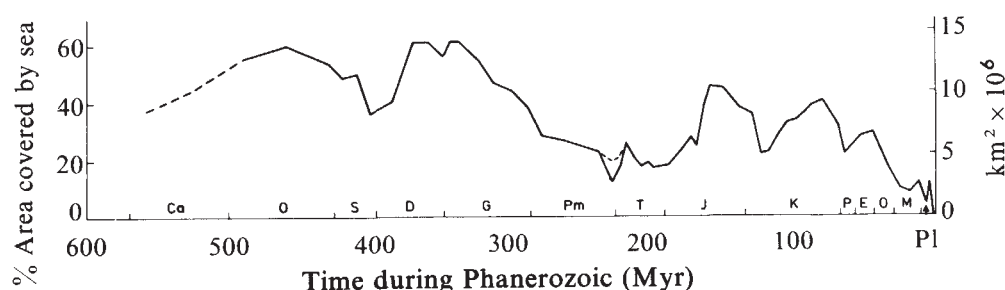
USSR by middle Ordovician times. Thereafter there was a faster withdrawal to reach a minimum value at the Silurian-Devonian boundary, followed by a relatively rapid restoration to middle Ordovician values in the late Devonian and early Carboniferous. Subsequently a progressive withdrawal culminated in a minimum at the Permian-Triassic boundary. During the Mesozoic, there were two transgressive peaks, in the late Jurassic and late Cretaceous, interrupted by an early Cretaceous minimum and followed by a sharp Cainozoic decline interrupted by a small peak in the middle to late Eocene, and a further, more dubious, peak in the Pliocene. None of the Mesozoic and Cainozoic peaks attained the values of the Palaeozoic maxima and in fact barely extended beyond the Silurian-Devonian trough.

The close spacing of data points, back at least to the Devonian, rules out a systematic temporal bias, with the length of time intervals chosen increasing the probability of overestimating the extent of sea the earlier the period⁹. Indeed, the probability of losing the stratigraphic record through subsequent deep burial, metamorphism and erosion must increase with time, and so the maps for older periods are more likely to underestimate than overestimate the former extent of marine cover. Furthermore, cratonic areas such as the Russian Platform and margins of the Siberian Shield, in fact almost everywhere outside eugeo-synclinal regions, contain much more substantial proportions of carbonates and evaporites to terrigenous clastics in the Palaeozoic than Mesozoic or Cainozoic, implying more areally restricted and topographically subdued sediment sources (mechanical denudation rates on the present continents show a tendency towards an exponential increase with increasing topographic elevation¹²). Hence the Russian data support strongly the notion of Phanerozoic secular withdrawal of sea.

North America

The only comprehensive series of palaeogeographic maps available for North America is that published in the atlas of Schuchert¹³, which was used by Wise⁹ to demonstrate a condition of essentially constant continental freeboard throughout the Phanerozoic, interrupted by short-term oscillations of sea level, 80% of which remained within about 60 m of a normal freeboard level $\sim 20 \text{ m}$ above the present level. These maps were drawn, however, several decades before the atlas was published, and therefore do not take

Fig. 1 Area of the Soviet Union covered by sea at different times during the Phanerozoic, expressed in both relative and absolute terms. Dating based on Geological Society of London Phanerozoic time scale. Ca, Cambrian; O, Ordovician; S, Silurian; D, Devonian; C, Carboniferous; Pm, Permian; T, Triassic; J, Jurassic; K, Cretaceous; P, Palaeocene; E, Eocene; O, Oligocene; M, Miocene; Pl, Plio-Pleistocene. Broken line signifies extent of non-marine Permo-Triassic.



into account extensive modern discoveries of strata, especially in the Arctic. Moreover, Schuchert left large areas of the cordilleran region blank for the Palaeozoic, but facies changes from areas further east clearly indicate a marine eugeosynclinal regime¹⁴. As with the USSR, Palaeozoic rocks of the North American central craton are much richer in carbonates than younger rocks, and testify to a correspondingly greater spread of sea, so that Schuchert's maps, based on erosional remnants, tend to be cautious and substantial underestimates for times earlier than the Mesozoic.

Data from a new atlas of facies distribution through time¹⁵ are here used to propose a revised graphical interpretation of North American data. In the absence of up-to-date palaeogeographic maps for individual stages, an attempt was made to determine the approximate maximum extent of sea for each period, using references cited in the atlas and making fairly conservative inferences. This allows a general comparison with Schuchert's data and the world data of the Termiers and Strakhov, plotted in a similar way (Fig. 2).

Schuchert considerably underestimated the former extent of seas from the Cambrian to the Lower Carboniferous. Subsequently, the Schuchert and revised North American curves match quite closely with only a relatively slight but systematic increase in inferred extent of flooding in the revised version. The revised curve matches the similarly plotted Russian curve much more closely than the Schuchert curve. The main differences are that the North American seas attained their maximum in the late Ordovician rather than the middle Ordovician and late Devonian-early Carboniferous, although the values are remarkably similar at around 60%, and the late Jurassic transgression was less extensive than the late Cretaceous. Because the two continents comprise approximately one-third of the Earth's land surface it might reasonably be inferred that the changes in question are primarily expressions of eustatic changes in sea level, with the relatively minor differences being attributable to regional differences in topography and rates of subsidence. This is confirmed by the close match of the two curves with those based on the Termier and Strakhov world palaeogeographic maps. A eustatic interpretation is both simpler and more plausible than the alternative proposed by Sloss and Speed¹⁶, that the continents have moved up and down in concert. An additional point is that the Permo-Triassic minimum must be due to a genuine eustatic fall and not to the silting up of shallow sea during a phase of stillstand. This is because the extent of end Permian and beginning Triassic non-marine sediments is negligible in both the USSR (Fig. 1) and North America.

Discussion

Several different explanations can account for the long-term secular changes under consideration. Earth expansion: besides the fact that the notion of slow earth expansion poses serious physical problems¹⁷ it has already been pointed out that the Phanerozoic overall regression can be explained more conservatively without considering a change in the earth's radius. In addition, the regression ought to have continued back into the Proterozoic, but the end of the Proterozoic was in fact a regressive phase. Thus Lower Cambrian deposits transgress widely on to much older rocks, often Archaean or early Proterozoic in age, a fact that has given rise to the term Lipalian interval. This regressive interval is probably comparable in magnitude with the end-Palaeozoic and late Cainozoic examples.

The alternative expansion model, which involves a much more rapid increase in the earth's radius since the early Mesozoic, leading to dispersal of the fragments of Pangaea^{6,18}, is even more clearly refuted by physical arguments¹⁹⁻²¹ and is incompatible with the dominantly

transgressive character of Mesozoic seas. Holmes⁴ even interpreted the Quaternary secular fall in sea level, superimposed on glacially-controlled oscillations, as a consequence of rapid earth expansion.

Glacial control: the onset of the Cainozoic ice age might be considered to have been responsible for a substantial component of the contemporary drop in sea level. It has been estimated that sea level would rise 46 m if all polar ice was melted⁷. An Antarctic ice sheet first formed in the late Miocene, although ice apparently started to accumulate as early as the late Eocene²². Abstraction of ice from the ocean system can only account for a small proportion of the lowering of sea level since the late Cretaceous high stand, however, if we accept the figure of slightly over 500 m computed from estimated changes through time of oceanic ridge volume²³, and supported independently by an estimate of the overall continental margin sedimentary offlap through the Cainozoic (P. R. Vail, personal communication).

The glacial explanation is even less successful in accounting for the major late Permian-early Triassic regression, which took place after the melting of the extensive late Carboniferous-early Permian Gondwana ice cap and the concomitant establishment of a more equable world climate.

Orogeny: plate tectonics provides a ready mechanism for the thickening and hence uplift of elongate sectors of crust at continental margins overlying subduction zones, by frictional melting of the descending slab giving rise to ascending magmas⁸, and perhaps more locally by the underthrusting of one continental edge beneath another. Thus closure of the Iapetus Ocean at the end of the Lower Palaeozoic in north-west Europe and eastern North America has converted a marginal basin-island arc-trench regime to a tectonically more stable zone of continent which has not been the site of subsequent orogeny but has persisted in most places as a positive area welded to the adjacent Precambrian shields and only locally and occasionally inundated by shallow sea.

With the detailed Russian data of Fig. 1, the clearest indication of orogeny correlating with regression concerns the substantial post-Valanginian eastward withdrawal of the sea in eastern Siberia, which is bound up with the younger Cimmerian Orogeny²⁴. The subsequent Aptian to Campanian transgression never reached as far west as the late Jurassic seas, and the Cimmerian event is the principal reason for the Russian late Cretaceous extent of sea being less than the late Jurassic, hence contradicting the world picture.

Similarly, the late Carboniferous-Permian regression is associated with the multiphase Hercynian Orogeny, commencing in the middle Carboniferous and reaching a climax in the early Permian, and converting the old Uralian and Angara geosynclines into cratonic regimes. More tentatively, the regression commencing in the late Ordovician and reaching a maximum at the end of the Silurian seems to be connected with the Russian equivalent of the Caledonian Orogeny, as manifested around the western and southern margins of the Siberian Shield. In the far east of the country, a notable end-Cretaceous orogeny is reported locally²⁴, and a number of minor Tertiary orogenic phases in the same region might help to account for part of the Cainozoic overall regression.

The North American data of Fig. 2 are, of course, less precise, but it is apparent that a component of the post-middle Ordovician Palaeozoic regression may be attributed to the progressive conversion of the Appalachian and neighbouring regions by the successive Taconic, Acadian and Appalachian orogenies from a geosynclinal-island arc zone into a positive upland area welded to the central North American craton, which lay beyond the reach of subsequent Mesozoic and Cainozoic transgressions. Like-

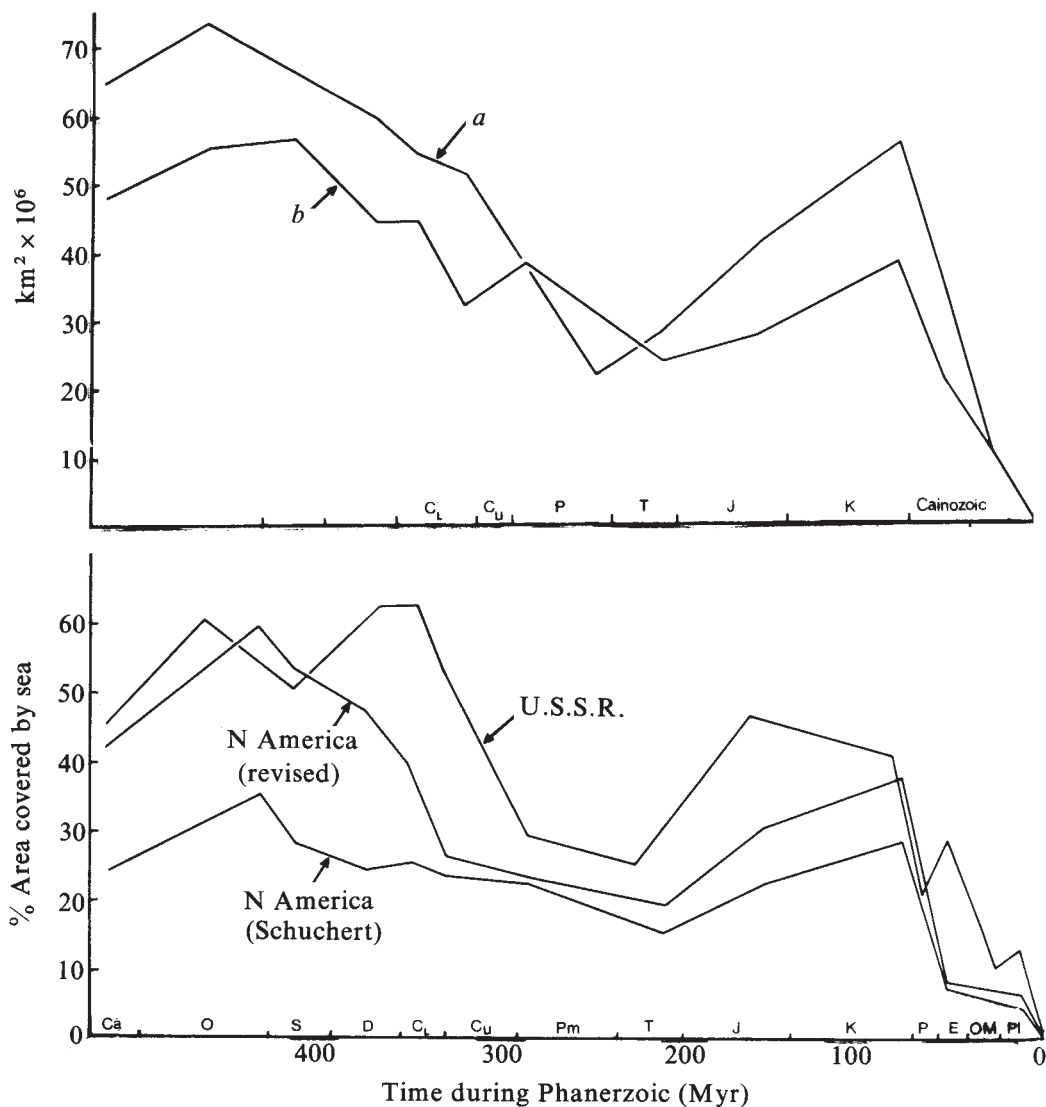


Fig. 2 Maximum degree of marine inundation for each geological period for USSR and North America (according to Schuchert and as revised in this paper, and comparison with the world data of Strakhov (a) and the Termiers (b), derived from ref. 8.

wise, the pronounced post-Cretaceous regression correlates with the Laramide Orogeny in the west of the continent.

That orogeny cannot account for all the regression is shown by examination of the history of extensive shield and platform areas of both continental masses, which were unaffected by Phanerozoic diastrophism. Thus the Russian Platform was extensively covered by shallow sea throughout the great bulk of Phanerozoic time but was completely emergent by the end of the Permian. A post-Cretaceous emergence is also apparent after the extensive late Mesozoic transgressions. Similarly, the Siberian Shield was substantially transgressed in the Cambro-Ordovician but a major withdrawal of the sea commenced in the west towards the close of the Silurian, with a land area emerging between the Ural and east Siberian seas. From middle Jurassic times onwards the sea covered the huge West Siberian depression but withdrew definitively in the Oligocene.

In North America, shallow seas covered a large part of the Canadian Shield from the Ordovician to the Devonian, but withdrew definitively in the Lower Carboniferous, as they did in the American mid-west in the Upper Carboniferous and Permian, and in the USA and Canadian Western Interior after the Cretaceous.

More generally, only a small percentage of the present land surface is occupied by 'Alpine' orogenic zones, yet it has been estimated that continental relief has increased by a factor of about 2.5 in the late Cainozoic²⁵.

Continental underplating: to account for the apparent emergence through the Phanerozoic of such cratonic areas

I formerly argued for a slight thickening by sialic underplating, using a term proposed for the African cratons by Shackleton²⁶. Several lines of evidence suggest, however, that old cratons have remained tectonically inert after the early Proterozoic and have been subjected to relatively little uplift or erosion since their time of formation, an interpretation supported by low heat flow at present²⁷. Persistence of ancient structural patterns and distinctive types of mineral assemblage is also incompatible with significant change at the base of the crust²⁸. Underplating, in fact, invokes a petrological process not required by plate tectonics and not supported by independent evidence. It is probably only relevant, if at all, to the Archaean.

Variations in rate of seafloor spreading: Hays and Pitman²³ attribute the late Cretaceous major transgression to a rapid seafloor spreading in the Pacific and Atlantic from 110 to 85 Myr, with a correspondingly greater extent of 'hot', buoyant oceanic ridge displacing seawater on to the continents. This interpretation relies on the accuracy of correlation of magnetic anomalies and has been questioned by Berggren *et al.*³⁰, who proposed an amended time scale consistent with more-or-less uniform rate of spreading. Even disregarding this criticism, the Hays and Pitman interpretation does not account satisfactorily for the marked secular regression through the Cainozoic. According to the data presented by Larsen and Pitman²⁹, the spreading rate for the Atlantic remained more or less constant from Palaeocene time onwards, and matched that for the late Jurassic and early Cretaceous. The rate was similarly con-

stant for the Phoenix–Pacific plate boundary, while the fuller data for the Farallon–Pacific plate boundary shows a fall at 85 Myr to a lower rate from the Coniacian to the Palaeocene. Thereafter there was a sharp but small reduction at the Palaeocene–Eocene boundary, a rise to a maximum in the Oligocene and subsequently a stepwise decline to the present. The Oligocene, however, is clearly regressive with respect to the Palaeocene and Eocene³¹.

Taking Indian Ocean data into account only complicates the picture further, and in a way not obviously consistent with the Hays and Pitman interpretation³². India moved away from Antarctica at a very rapid rate in the early Tertiary, and was followed by a period when little or no spreading took place west of the 90°E Ridge. Spreading from the Carlsberg Ridge has proceeded at a more or less constant rate of 1.2–1.3 cm yr⁻¹ over the past 30–35 Myr, back into the Oligocene, following a period of very slow spreading back to the late Palaeocene. In the central Indian Ocean the spreading rate has not been constant as in the Atlantic and Pacific, but changed from 5.7 cm yr⁻¹ in the Campanian to ~9 cm yr⁻¹, persisted at that rate until the middle Palaeocene and then decelerated.

Changes in length of ocean ridge system: the most plausible alternative to the spreading rate interpretation to account for the spectacular late Cretaceous transgression, which was obviously a very unusual geological event, is that a large volume of seawater was displaced by creation of new oceanic ridges in the South Atlantic and Indian oceans, since this was the time when effective dispersal of the fragments of Gondwana commenced. Such an interpretation would imply that the Cainozoic regression was due at least partly to the progressive loss of ridges elsewhere by subduction. That this is likely to have happened is evident from the analysis of Larsen and Pitman²⁸. According to their interpretation, in the middle Cretaceous there were four major plates in the Pacific region, the Kula, Farallon, Phoenix and Pacific plates. In the past 110 Myr there has been progressive consumption through subduction beneath Asia and the Americas of ridges separating the Kula from the Farallon, and the Farallon from the Phoenix plate, together with the Pacific–Kula ridge.

On the other hand, the early Tertiary saw the creation of new ridges between Greenland and Scandinavia as the North Atlantic opened, and between Antarctica and Australia. This might account for the worldwide Eocene transgression³¹ shown for the USSR in Fig. 1, which is against the secular Cainozoic trend.

Assessment of the role of ridges in pre-Cretaceous times must be highly speculative because the marine record is lost, but it seems possible that the more or less progressive Jurassic transgression relates in part to the creation of a ridge within the newly-opened central Atlantic, while the striking late Palaeozoic regression could to some extent be due to the consumption of spreading sea floor including ridges that had previously driven continents together to effect the Hercynian Orogeny.

The dominant cause of the longer-term Phanerozoic changes in sea level and overall regression is likely to have been a combination of continental thickening by orogeny consequent on subduction or collision, and variation in the cumulative length of the oceanic ridge system. The best prospect for testing these ideas quantitatively is in the Cainozoic, which has by far the fullest record of geological events per unit of time, both on the continents and ocean floor.

Received 28 June; accepted 2 September 1977.

1. Termier, H. & G. *Histoire Géologique de la Biosphère* (Masson, Paris, 1952).
2. Strakhov, N. M. *Outlines of Historical Geology* (Government Printing Office, Moscow, 1948).
3. Eged, L. *Geophys. Pura Apl.* 33, 42 (1956).
4. Holmes, A. *Principles of Physical Geology* 2nd edn (Nelson, London, 1965).
5. Dearnley, R. *Physics and Chemistry of the Earth* (eds Ahrens, L. H. et al.) 7, 1 (1966).
6. Carey, S. W. *Earth Sci. Rev.* 11, 105 (1975).
7. Armstrong, R. L. *Nature* 221, 1042 (1969).
8. Hallam, A. *Nature* 232, 180 (1971).
9. Wise, D. U. *The Geology of Continental Margins* (eds Burk, C. A. & Drake, C. L.) 45 (Springer, New York, 1974).
10. Vinogradov, A. P. (ed.) *Atlas of the Lithological–Palaeogeographic Maps of the USSR* 4 vols (Ministry of Geol., Moscow, 1967–69).
11. Hallam, A. *Earth Sci. Rev.* 5, 45 (1969).
12. Garrels, R. M. & Mackenzie, F. T. *Evolution of Sedimentary Rocks* (Norton, New York, 1971).
13. Schuchert, C., *Atlas of Paleogeographic Maps of North America* (Wiley, New York, 1955).
14. Johnson, J. G. *Bull. geol. Soc. Am.* 82, 3263 (1971).
15. Cook, T. D. & Bally, A. W. *Stratigraphic Atlas of North and Central America* (Princeton University Press, Princeton, 1977).
16. Sloss, L. L. & Speed, R. C. S.E.P.M. Spec. Publ. 22, 89 (1974).
17. Birch, F. *Phys. Earth planet. Inter.* 1, 141 (1968).
18. Owen, H. G. *Phil. Trans. R. Soc. A* 281, 223 (1976).
19. Hospers, J. & Van Andel, S. I. *Tectonophysics* 5, 5 (1967).
20. Le Pichon, X. *J. geophys. Res.* 73, 3661 (1968).
21. Stewart, A. D. *J. geol. Soc.* 133, 281 (1977).
22. Hayes, D. E. & Frakes, L. A. *Init. Rep. D.S.D.P.* 28, 919 (1975).
23. Hays, J. D. & Pitman, W. C. *Nature* 246, 18 (1973).
24. Nalivkin, D. V. *Geology of the U.S.S.R.* (Oliver and Boyd, Edinburgh, 1973).
25. Flint, R. F. *Glacial and Pleistocene geology* (Wiley, New York, 1957).
26. Shackleton, R. M. *Proc. geol. Ass. Lond.* 81, 549 (1970).
27. Watson, J. V. *Phil. Trans. R. Soc. A* 280, 629 (1976).
28. Watson, J. V. *Phil. Trans. R. Soc. A* 273, 443 (1973).
29. Larsen, R. L. & Pitman, W. C. *Bull. geol. Soc. Am.* 83, 3645 (1972).
30. Berggren, W. A., Mckenzie, D. P., Slater, J. G. & Van Hinte, J. E. *Bull. geol. Soc. Am.* 86, 267 (1975).
31. Hallam, A. *Am. J. Sci.* 261, 397 (1963).
32. McKenzie, D. P. & Slater, J. G. *Geophys. J.* 25, 437 (1971).

Location and bond type of intermolecular contacts in the polymerisation of haemoglobin S

Ruth E. Benesch, Suzanna Kwong, Reinhold Benesch & Rohinton Edalji

Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York, New York 10032

The solubility of 14 hybrid haemoglobins composed of α chains with a single substitution and β chains from HbS was compared with that of sickle haemoglobin. A substantial reduction in the insolubility of native deoxyhaemoglobin S results from surface mutations in certain regions of the α chain while changes in other areas have no effect. Also, the chemical nature of the substitution is decisive and points to the type of intermolecular bonding at several loci.

WE described previously the preparation of five haemoglobin double mutants carrying the $\beta 6$ Glu $\rightarrow$ Val mutation as well as an additional one on the α chains^{1,2}. They were used to delineate regions on the surface of the molecule where significant bonding between neighbouring tetramers takes place in the polymerisation of deoxyhaemoglobin S. This work has now been extended to include portions of the α chain which were previously unexplored. In addition, the availability of more than one mutation at the same locus has now made it possible to infer the type of bonding involved in some of these intermolecular contacts.

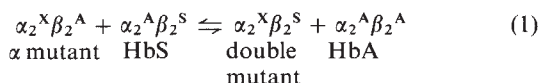
Table 1 Mutant haemoglobins used for preparation of HbS hybrids

No.	Mutant Hb	Locus on α chain	Substitution	Ref.	Method of isolation*	
					α Mutant	Hybrid
1	Sawara	6	Asp→Ala	6	6	1
2	Anantharaj	11	Lys→Glu	7	4	2
3	HbI	16	Lys→Glu	8	3	5
4	Sealy (Hasharon)	47	Asp→His	9,10	2	1
5	Mugino (Kokura)	47	Asp→Gly	11	1	1
6	Montgomery	48	Leu→Arg	12	1	1
7	Shimonoseki	54	Gln→Arg	13	2	1
8	J Mexico	54	Gln→Glu	14	4	2
9	G Phila	68	Asn→Lys	15	—	1
10	Ube II	68	Asn→Asp	16	4	2
11	Mahidol	74	Asp→His	17	6	1
12	Winnipeg	75	Asp→Tyr	18	6	1
13	Q Iran	75	Asp→His	19	6	1
14	O Indonesia	116	Glu→Lys	20	2	1

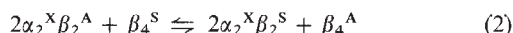
*Conditions for isolation of α mutant and hybrid haemoglobin were as follows: 1, Chromatography on DEAE-Sephadex A-50 and elution with a 24-h linear gradient of 0.05 M Tris-HCl, pH 8.4→7.4. 2, As(1), but pH 8.0→7.4. 3, As(1) but 12-h gradient pH 8.0→7.0. 4, As(1) but 16-h gradient pH 7.7→7.0. 5, As(1) but pH 7.6→6.8. 6, Polyacrylamide gel electrophoresis²¹.

Preparation and characterisation of double mutant haemoglobins

Two different hybridisation reactions were used to prepare double mutant tetramers. The first^{1,2} is based on the exchange reaction



This method is suitable for separating the double mutant except where the mutation on the α chain brings about a change in charge which is equal and opposite to that of the β^S mutation, as for example in HbJ Mexico ($\alpha^{54\text{Gln} \rightarrow \text{Glu}}$). In such a case, the hybrid, $\alpha_2^x\beta_2^S$, will, of course, have the same charge as HbA and can therefore not be separated from it. This difficulty can be overcome by hybridisation of such α mutants with β_4^S since HbA is not a product of this reaction



Method (2) was therefore used in all cases where the effective charge of the double mutant was too near to that of HbA to permit a good separation. This was often encountered in haemoglobins where the formal difference in charge brought about by the mutation was modified by interactions with other residues.

HbS was isolated from the blood of heterozygous donors by chromatography on DEAE-Sephadex³ using a 24-h linear gradient of 0.05 M Tris buffer pH 8.0–7.4.

Hb β_4^S was prepared by treatment of the unfractionated A/S haemolysate with *p*-mercuribenzoate followed by chromatography on CM-cellulose as described by Bucci and Fronticelli⁴. The β^{SH} fraction was collected, concentrated by ultrafiltration using PM 30 Amicon membranes and the mercury was removed with *N*-acetylpenicillamine as described previously⁵.

The α mutant haemoglobins including the five which were hybridised with HbS before^{1,2}, are listed in Table 1. Only mutants were chosen with substitutions at or near the surface of the molecule. The exact position of each residue was located on a three-dimensional model assembled from Labquip components using the refined coordinates of human deoxyhaemoglobin²². The haemoglobin variants were isolated from the blood of heterozygous donors. The samples were transported in temperature controlled containers (Trans-Temp 313, Fisher Scientific). The purity of the mutant haemoglobins was checked by polyacrylamide gel electrophoresis.

Hybrid tetramers with two β^S chains and α chains from mutants 3, 4, 5, 6, 7, 9 and 14 (Table 1) were prepared by method (1) as described previously^{1,2}. The remainder were made by method (2) using β_4^S as donor of β^S chains. Equal amounts of 1% solutions of α mutant haemoglobin and β_4^S were mixed and incubated in 0.2 M

acetate buffer pH 4.7 for 3–4 h at 0 °C. The $\alpha_2^{\text{Shimo}}\beta_2^S$ hybrid was made by both methods and HbS was also prepared from β_4^S and HbA. The conditions for the isolation of the double mutant haemoglobins are given in Table 1. The identity and purity of the hybrids was established by polyacrylamide gel electrophoresis of the intact haemoglobins and, in some cases, of the mercurocurated subunits.

The oxygen equilibrium curves of all haemoglobins including the hybrids were measured as described before^{23,24} using 10 μ M haemoglobin in 0.05 M bis-Tris buffer pH 7.4 at 20 °C.

The structural and functional integrity of the hybrid tetramers was tested using the criteria applied previously^{1,2}. The oxygen affinities of all the double hybrids were the same as those of the α mutant from which they were derived and, with two exceptions, were within the normal range ($\log p_{50} = 0.43$ – 0.56 at 20 °C and pH 7.4). Hb Mugino and $\alpha_2^{\text{Mugino}}\beta_2^S$ had a somewhat raised oxygen affinity ($\log p_{50} = 0.36$) and Hb Sawara was found to be a really high affinity mutant ($\log p_{50} = 0.02$), contrary to a previous report¹¹. In this case too, the Sawara-S hybrid retained the high affinity.

The oxygenation properties of HbS prepared from HbA and β_4^S by method (2) were identical with those of native HbS ($\log p_{50} = 0.47$).

Solubility in strong phosphate as measure of HbS polymerisation

The solubility in concentrated phosphate buffer, as originally described by Itano²⁵, was chosen for comparing the polymerising tendency of the double mutants with that of HbS, since only small quantities of the double mutants were available (15–30 mg). Several lines of evidence support the assumption that this method measures the same phenomenon as the usual gelling tests on much larger samples: (1) While the precipitation of oxy and deoxyhaemoglobin A as well as that of oxyhaemoglobin S is essentially instantaneous when they are mixed with the phosphate buffer, deoxyhaemoglobin S precipitates slowly, which presumably reflects the time required for the formation of ordered polymers. The suspensions were therefore incubated with gentle shaking for at least 30 min at 25 °C and the filtrates were observed for further precipitation. Incomplete precipitation introduces increasingly large errors as the fraction of haemoglobin in solution decreases with increasing ionic strength. (2) Cottam and Waterman²⁶ concluded from the parallel temperature dependence of the solubility and the gelation that the former constitutes 'a sensitive method of monitoring the aggregation phenomenon of deoxyhaemoglobin S'. (3) Both with some chemically modified HbS molecules²⁷ and in the case of two naturally occurring double mutants, the results obtained by the two methods were similar. Thus HbC Harlem was 2.5 times more soluble than HbS in strong phosphate²⁸ and its minimum gelling concentration was 36 g % instead of 24.5 g % for HbS (ref. 28). Hassan *et al.*²⁹ have shown

Table 2 Solubility of deoxygenated double mutants

Haemoglobin	α -chain locus	Solubility (mol l ⁻¹ × 10 ⁵)
$\alpha_2^A\beta_2^S$	—	0.3
$\alpha_2^{\text{Winnipeg}}\beta_2^S$	75	0.2
$\alpha_2^{\text{Sawara}}\beta_2^S$	6	0.2
$\alpha_2^{\text{Q Iran}}\beta_2^S$	75	0.3
$\alpha_2^{\text{Mahidol}}\beta_2^S$	74	0.3
$\alpha_2^{\text{I}}\beta_2^S$	16	0.9
$\alpha_2^{\text{Ube II}}\beta_2^S$	68	0.9
$\alpha_2^{\text{Mugino}}\beta_2^S$	47	0.9
$\alpha_2^{\text{G Phila}}\beta_2^S$	68	1.1
$\alpha_2^{\text{Anantharaj}}\beta_2^S$	11	1.5
$\alpha_2^{\text{Montgomery}}\beta_2^S$	48	1.5
$\alpha_2^{\text{J Mexico}}\beta_2^S$	54	1.7
$\alpha_2^{\text{Shimo}}\beta_2^S$	54	1.8
$\alpha_2^{\text{O Indonesia}}\beta_2^S$	116	2.0
$\alpha_2^{\text{Sealy}}\beta_2^S$	47	4.6

The measurements were carried out as described previously³⁰ in 2.04 M phosphate buffer pH 6.8 ($\mu = 4.4$) at 25 °C. The values were obtained by interpolation or extrapolation of solubility plots such as those shown in Fig. 1.

that the double mutant, Hb Stanleyville II/S ($\alpha_2^{78 \text{ Lys}}\beta_2^{6 \text{ Val}}$) is twice as soluble in strong phosphate as native HbS and this is paralleled by an increase in the minimum gelling concentration from 24 to 30 g %.

Effect of α -chain mutations on the solubility of HbS

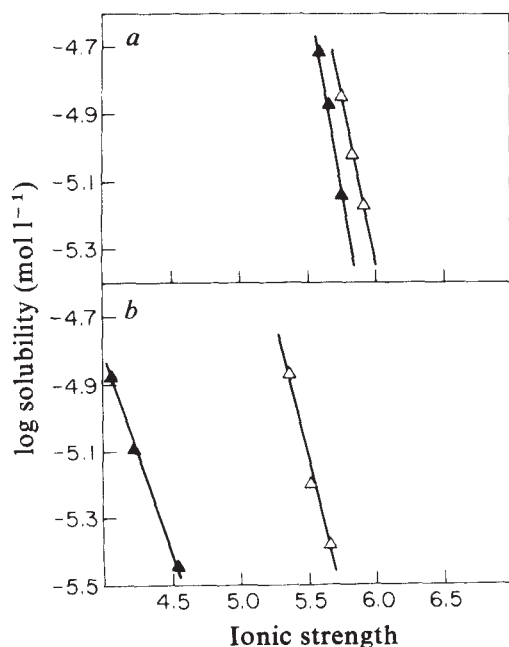
The solubility of the deoxy form of all the double mutants used in 4.4 ionic strength phosphate buffer is given in Table 2. Clearly, the effect of the mutations varies over a wide range from no effect at all at positions 6, 74 and 75 to the 20-fold increase in solubility caused by the replacement of aspartic acid by histidine at position 47.

Another criterion was also used to compare the various double mutants with HbS—the change in ionic strength ($\Delta\mu$) necessary to make the solubility of the oxy and deoxy form equal in each case (Fig. 1). This value was 1.35 for HbS and 0.22 for HbA.

The effect of a mutation in the α chain on the solubility of the tetramer it forms with β^S chains depends not only on the location of the substitution but also on its chemical nature.

The most instructive from this point of view were the three pairs of mutants at positions $\alpha 47$, $\alpha 54$ and $\alpha 68$. The results obtained on the hybrids of these mutants with HbS are shown in Table 3. The

Fig. 1 Solubility of haemoglobins A (a) and S (b) in phosphate buffer pH 6.8 at 25 °C. Open symbols, oxyhaemoglobin; filled symbols, deoxyhaemoglobin.

**Table 3** Effect of α -chain mutations on solubility of HbS

Haemoglobin	Position on α chain	Residue	$\Delta\mu^*$
$\alpha_2^A\beta_2^S$	47 (CD6)	Asp	1.35
$\alpha_2^{\text{Mugino}}\beta_2^S$		Gly	1.07
$\alpha_2^{\text{Sealy}}\beta_2^S$		His	0.58
$\alpha_2^A\beta_2^S$	54 (E3)	Gln	1.35
$\alpha_2^{\text{Shimo}}\beta_2^S$		Arg	0.78
$\alpha_2^{\text{J Mexico}}\beta_2^S$		Glu	0.78
$\alpha_2^A\beta_2^S$	68 (E17)	Asn	1.35
$\alpha_2^{\text{G Phila}}\beta_2^S$		Lys	1.08
$\alpha_2^{\text{Ube II}}\beta_2^S$		Asp	1.10

*Change in ionic strength necessary to make the deoxy and oxy forms equally soluble.

greatest solubilising effect on deoxyhaemoglobin S encountered among 14 different substitutions on the α chain is that due to replacement of aspartic acid by histidine at $\alpha 47$ (Hb Sealy). But, when the substitution at this locus is glycine, as in Hb Mugino, the solubilising influence is greatly reduced (Table 3). We therefore conclude that the aspartate side chain at $\alpha 47$ is engaged in an intermolecular salt bridge which materially contributes to the stability of the helical polymer. Wishner *et al.*³¹ and Edelstein *et al.*³² also concluded that $\alpha 47$ is in an important contact region. If the spacial relations in the fibre can be assumed to be similar to the arrangement of the double strands seen in crystals of deoxyhaemoglobin S the most likely partner for aspartate $\alpha 47$ is lysine $\beta 59$ on the adjacent molecule³¹.

Table 3 shows that the glutamine residue at $\alpha 54$ also provides an important intermolecular contact point. Here, however, the same increase in solubility is observed whether the glutamine is replaced by a positive residue (Hb Shimonoseki) or a negative residue (HbJ Mexico). Although the solubility data show that $\alpha 68$ is a weaker contact point altogether, the nature of the bond must be of the same kind as that at $\alpha 54$, since here again, both a negative and a positive substitution have the same effect. This evidence suggests apolar bonds at these two locations. Hydrophobic bonding is generally assumed to be important in the polymerisation of HbS as, for example, indicated by the temperature dependence of the gel formation³³. In this connection it is pertinent to recall a parallel situation involving the β chains where no gelation occurs when the valine at $\beta 6$ is replaced either by glutamic acid (HbA) or by lysine (HbC).

A second region of the tetramer surface formed by the α chains qualifies as a weaker contact area. It contains residues $\alpha 11$, $\alpha 16$ and $\alpha 116$, brought into proximity by the tertiary structure of the protein. The mutations at these loci (Hb Anantharaj, HbI and HbO Indonesia) all have comparable effects on the solubility of HbS (Table 4).

A particularly significant group of mutants were Hb Sawara ($\alpha 6$), Hb Mahidol ($\alpha 74$) and haemoglobins Winnipeg and Q Iran ($\alpha 75$). In all these cases the hybrid composed of the abnormal α chains and β^S chains had the same solubility as HbS itself (Table 5). The three-dimensional model of the tetramer shows that these three loci ($\alpha 6$, $\alpha 74$ and $\alpha 75$) are situated quite close together with $\alpha 6$ being somewhat recessed from the surface. This part of the surface can therefore be designated as a non-contact region. It should be noted that two of these mutants—haemoglobins Mahidol and Q Iran—carry the same substitution (Asp→His) as does Hb Sealy

Table 4 Effect of α -chain mutations on solubility of HbS

Haemoglobin	Position on α chain	Substitution	$\Delta\mu^*$
S	—	—	1.35
$\alpha_2^{\text{Anantharaj}}\beta_2^S$	11 (A9)	Lys→Glu	0.86
$\alpha_2^{\text{I}}\beta_2^S$	16 (A14)	Lys→Glu	1.15
$\alpha_2^{\text{O Indonesia}}\beta_2^S$	116 (GH4)	Glu→Lys	0.84

*Change in ionic strength necessary to make the deoxy and oxy forms equally soluble.

Table 5 Effect of α -chain mutations on solubility of HbS

Haemoglobin	Position on α chain	Substitution	$\Delta\mu^*$
S			1.35
α_2 Sawara β_2^S	6 (A4)	Asp→Ala	1.37
α_2 Mahidol β_2^S	74 (EF3)	Asp→His	1.35
α_2 Winnipeg β_2^S	75 (EF4)	Asp→Tyr	1.33
α_2 Q Iran β_2^S	75 (EF4)	Asp→His	1.33

*Change in ionic strength necessary to make the deoxy and oxy forms equally soluble.

($\alpha 47$). As the same change has a maximal effect at $\alpha 47$, but none at $\alpha 74$ and $\alpha 75$, the location on the surface of the molecule is clearly the decisive factor.

The Hb Sawara hybrid also shows that a high oxygen affinity is not incompatible with efficient polymerisation of deoxyhaemoglobin S. The same situation was encountered previously with HbS from which the α C-terminal residue had been removed—desArg HbS (ref. 1).

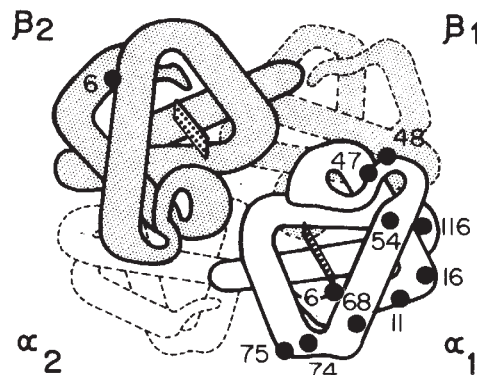
This method of approach has thus made it possible to define several distinct areas on the surface of the haemoglobin molecule formed by the α chains. One of these is a region of strong intermolecular bonding which includes aspartate $\alpha 47$, leucine $\alpha 48$ and glutamine $\alpha 54$. Of these, $\alpha 47$ aspartate probably forms a salt link with lysine $\beta 59$ on the next molecule while glutamine $\alpha 54$ participates in hydrophobic bonding.

A region where replacement of the normal residues has a smaller effect in solubilising HbS is enclosed by the charged residues at $\alpha 11$, $\alpha 16$ and $\alpha 116$. All three are therefore candidates for electrostatic bonding. Another contact point of comparable significance is the asparagine at $\alpha 68$, engaged in hydrophobic bonding. A weak effect brought about by a single substitution could be the result of multiple contacts in that region so that replacement of only one of them would weaken the contact by only a fractional amount.

Finally, the mutations at $\alpha 6$, $\alpha 74$ and $\alpha 75$ provide an excellent control, since changes at these positions—and in the case of $\alpha 75$ even two different substitutions—have no effect whatever on the solubility of native HbS.

In Fig. 2, the 10 different loci with one or more substitutions, which were included in this study are marked on a schematic representation of the haemoglobin tetramer. But, the distance between the different positions can only be judged from the three-dimensional model and not from this necessarily two-dimensional substitute.

We thank Drs S. Pootrakul, S. Rahbar, D. Labie, J. Rosa, J. Alperin, M. Grey, R. Rieder, T. Hosty, P. Galbraith, T. Imamura and T. Miyaji for blood samples. This work was supported by US PHS grants HL-17552 and HL-05791 from the National Heart and

**Fig. 2** Schematic model of the haemoglobin molecule showing the locations of the amino acid substitutions.

Lung Institute and grant PCM76 20361 from the NSF. R.B. is a recipient of a US PH research career award from the National Heart and Lung Institute.

Received 24 June; accepted 22 August 1977.

1. Benesch, R. E., Yung, S., Benesch, R., Mack, J. & Schneider, R. G. *Nature* **260**, 219–221 (1976).
2. Benesch, R. E., Yung, S., Benesch, R., Mack, J. & Schneider, R. G. *Proc. Symp. Molecular and Cellular Aspects of Sickle Cell Disease* Dallas, Texas, DHEW Publ. No. 76-1007, 113–131 (1975).
3. Huisman, T. H. J. & Dozy, A. M. *J. Chromat.* **19**, 160–169 (1965).
4. Bucci, E. & Fronticelli, C. *J. Biol. Chem.* **240**, PC551 (1965).
5. Benesch, R., Benesch, R. E. & Enoki, Y. *Proc. natn. Acad. Sci. U.S.A.* **61**, 1102–1106 (1968).
6. Sumida, I., Ohta, Y., Imamura, T. & Yanase, T. *Biochim. biophys. Acta* **322**, 23–26 (1973).
7. Pootrakul, S., Kematarn, B., Na-Nakorn, S. & Suanpan, S. *Biochim. biophys. Acta* **405**, 161–166 (1975).
8. Schneider, R. G., Alperin, J. B., Beale, D. & Lehmann, H. *J. lab. clin. Med.* **68**, 940–946 (1966).
9. Halbrecht, I., Isaacs, W. A., Lehmann, H. & Ben-Porat, F. *Israel J. med. Sci.* **3**, 827–831 (1967).
10. Schneider, R. G., Ueda, S., Alperin, J. B., Brimhall, B. & Jones, R. T. *Am. J. hum. Genet.* **20**, 151–156 (1968).
11. Sumida I. *Jap. J. hum. Genet.* **19**, 343–363 (1975).
12. Brimhall, B. *et al. Biochim. biophys. Acta* **379**, 28–32 (1975).
13. Miyaji, T., Iuchi, I., Takeda, I. & Shibata, S. *Acta Jaem. Jap.* **26**, 531–537 (1963).
14. Jones, R. T., Koler, R. D. & Lisker, R. *Clin. Res.* **11**, 105 (1963).
15. Baglioni, C. & Ingram, V. M. *Biochim. biophys. Acta* **48**, 253–265 (1961).
16. Miyaji, T., Iuchi, I., Yamamoto, K., Ohba, Y. & Shibata, S. *Clin. chim. Acta* **16**, 347–352 (1967).
17. Pootrakul, S. & Dixon, G. H. *Can. J. Biochem.* **48**, 1066–1078 (1970).
18. Vella, F., Wiltshire, B. G., Lehmann, H. & Galbraith, P. *Clin. Biochem.* **6**, 66–70 (1973).
19. Lorkin, P. A. *et al. Br. J. Haemat.* **19**, 117–125 (1970).
20. Baglioni, C. & Lehmann, H. *Nature* **196**, 229–232 (1962).
21. Suzuki, T., Benesch, R. E., Yung, S. & Benesch, R. *Analyt. Biochem.* **55**, 249–254 (1973).
22. Fermi, J. *molec. Biol.* **97**, 237–256 (1975).
23. Benesch, R., Macduff, G. & Benesch, R. E. *Analyt. Biochem.* **11**, 81–87 (1965).
24. Benesch, R. E., Benesch, R. & Yung, S. *Analyt. Biochem.* **55**, 245–248 (1973).
25. Itano, H. A. *Archs Biochem. Biophys.* **47**, 148–159 (1953).
26. Cottam, G. L. & Waterman, M. R. *Biochem. biophys. Res. Commun.* **54**, 1157–1163 (1973).
27. Benesch, R., Benesch, R. E. & Yung, S. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1504–1505 (1974).
28. Bookchin, R. M., Nagel, R. L. & Ranney, H. M. *J. mol. Chem.* **242**, 248–255 (1967).
29. Hassan, W., Beuzart, Y., North, M. L. & Rosa, J. *Biochem. biophys. Res. Commun.* (in the press).
30. Benesch, R., Benesch, R. E. & Yung, S. *Biochem. biophys. Res. Commun.* **55**, 261–265 (1973).
31. Wishner, B. C., Hanson, J. C., Ringle, W. M. & Love, W. E. *Proc. Symp. Molecular and Cellular Aspects of Sickle Cell Disease*, Dallas, Texas, DHEW Publ. No. 76-1007, 1–29 (1975).
32. Edlstein, S. J. *et al. Proc. Symp. Molecular and Cellular Aspects of Sickle Cell Disease*, Dallas, Texas, DHEW Publ. No. 76-1007, 33–59 (1975).
33. Allison, A. C. *Biochem. J.* **65**, 212–219 (1957).

Synchronised transmembrane insertion and glycosylation of a nascent membrane protein

James E. Rothman & Harvey F. Lodish

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Studies of the synthesis and incorporation of the vesicular stomatitis virus glycoprotein into membranes in a synchronised cell-free system demonstrate a tight coupling between polypeptide synthesis and membrane insertion, as a result of which the nascent chain crosses the membrane. The studies reveal a surprisingly precise sequence by which the nascent chain of this membrane glycoprotein is glycosylated in two steps. These findings have important implications for the mechanisms of membrane assembly.

MEMBRANES are continuously formed and remodelled in all living cells, and the problem of how membranes are assembled is central to cell biology. But, the mechanisms by which proteins are incorporated into membranes and the means by which specific proteins are directed only to the appropriate membrane or organelle are obscure. We describe here a novel approach to these problems which has begun to elucidate the nature and temporal sequence of the events responsible for the synthesis and membrane insertion of ectoproteins¹, those integral membrane proteins which have a hydrophilic domain on the extra-cytoplasmic side of the membrane. The external portion(s) of

ectoproteins must somehow cross the permeability barrier of the membrane at some stage. Several considerations¹ suggest that the external portions of the polypeptide chain of an ectoprotein cross the membrane during protein synthesis, but direct evidence has been lacking. The alternative school of thought² holds that ectoproteins can be inserted into membranes after they are completely synthesised.

The glycoprotein (G) of vesicular stomatitis virus (VSV) is such an ectoprotein. Most of the protein of G and all of its carbohydrate are external to the lipid bilayer of the viral envelope, forming the 'spikes' which project radially from the surface of the viral particle³. The G protein is synthesised by ribosomes bound to the endoplasmic reticulum^{4,5} of infected cells and receives at least the core *N*-acetylglucosamine and mannose sugars in this organelle⁶. G is a transmembrane protein of the endoplasmic reticulum membrane, oriented with its sugar-containing portion within the lumen of the organelle, but with about 30 amino acids at its carboxy terminus exposed on the opposite, cytoplasmic side of the membrane^{6,9}. G is ultimately transported from the reticulum to the plasma membrane, during which time the glycosylation is completed by the addition of terminal sialic acid residues^{6,7}, and is subsequently incorporated into the viral envelope following the budding out of the viral nucleocapsid through the cell surface membrane³.

A cell-free system capable of supporting membrane assembly has been described⁹ in which the VSV G protein was translated from viral-specific mRNA in crude extracts of wheat germ. When G was synthesised in the presence of membrane vesicles derived from the rough endoplasmic reticulum of pancreas (but freed of endogenous ribosomes), this membrane protein was glycosylated and was asymmetrically inserted into the vesicle membrane, spanning the lipid bilayer, with the same orientation as the native G protein found in the rough endoplasmic reticulum *in vivo*.

We report here a simple method for synchronising the growth and processing of the G protein in this cell-free system. This technique has revealed a remarkably close degree of coupling between polypeptide synthesis and incorporation into the membrane, and has made it possible to demonstrate a precise sequence by which the nascent chain of the membrane protein is glycosylated.

Synchronisation of protein synthesis

Protein synthesis is started when total VSV mRNA is added to a cell-free extract of wheat germ. After 2 min of protein synthesis, a specific inhibitor of initiation¹⁰, 7-methylguanosine-5'-phosphate (⁷mGp) is added to prevent subsequent initiations. If this interval is sufficiently short in comparison with the time required to complete the G protein, and if all of the nascent glycoprotein chains grow at the same rate, then in principle, all copies of the G protein will be polymerised in synchrony as a cohort population.

Figure 1 shows that an approximation to synchrony can be obtained by this procedure. The synthesis of total VSV protein proceeds linearly for a limited period (Fig. 1*a*), beyond which no further protein is made, as would be expected if initiation were completely blocked by ⁷mGp. In the absence of ⁷mGp, synthesis is linear for about 60 min (not shown). No completed G protein can be observed until about 45 min, when all chains of G are observed to be completed within a rather brief time span of 10 min (Fig. 1*b*). Total protein synthesis is apparently complete before the appearance of G protein because G is a minor fraction of the total translation products and greatly exceeds most of the products in chain length.

Coupling of insertion and glycosylation to protein synthesis

To determine the degree to which the processes of insertion and glycosylation of G are coupled to protein synthesis, synthesis of G chains was initiated synchronously in the absence of pancreatic membrane vesicles, and at various times samples were removed and the chains were completed in the presence of vesicles. This

experiment determines the maximum size of nascent glycoprotein which can be made without any association with membranes without affecting proper insertion and glycosylation.

As shown previously⁹, the partially glycosylated form of the glycoprotein (G₁) migrates more slowly (see Fig. 2) through sodium dodecyl sulphate-polyacrylamide gels than does the unglycosylated form (G₀) which results from synthesis in the absence of pancreatic membrane vesicles. The extent of glycosylation was therefore determined as the fraction of G

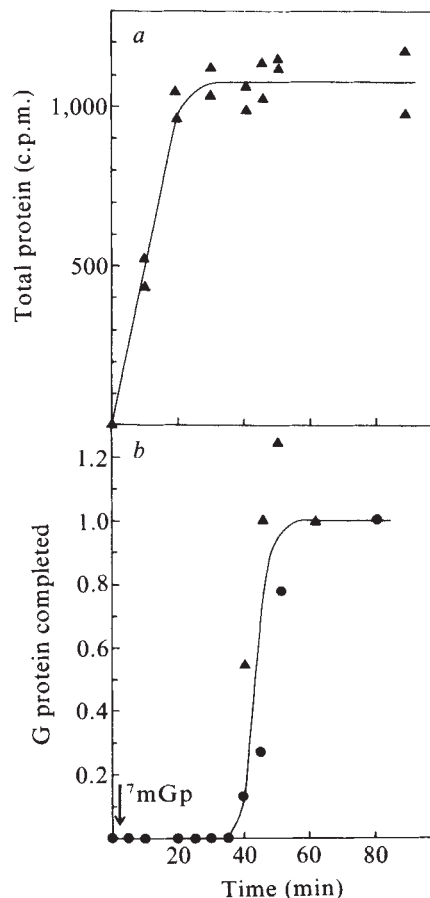


Fig. 1 Synchronisation of protein synthesis. Wheat germ extract and VSV were prepared and proteins were synthesised at 23 °C, essentially as before⁹, in a volume (including all additions) of 0.5 ml containing 0.15 ml wheat germ extract, 20 μ l VSV mRNA, 27 mM HEPES (pH 7.2), 120 mM KCl, 1.5 mM magnesium acetate, 5.6 mM dithiothreitol, 2.6 mM ATP, 0.26 mM GTP, 16 mM creatine phosphate, 0.73 mM spermidine, 42 μ g ml⁻¹ creatine phosphokinase, 0.13 mM each of the 19 amino acids, except methionine, and ³⁵S-methionine (0.5 mCi ml⁻¹) with a specific radioactivity of approximately 600 Ci mmol⁻¹. The mixture was pre-incubated for 1.5 min at 23 °C before adding mRNA at time zero. After 2 min, 25 μ l of a 20 mM solution of ⁷mGp (P. L. Biochemicals) were added, giving a final concentration of 1 mM. *a*, At various times, duplicate samples of 2 μ l were taken, and the total protein synthesised was determined as radioactivity insoluble in boiling 5% trichloroacetic acid (TCA). The amount of acid-insoluble radioactivity at zero time (550 c.p.m.) was subtracted to yield the data in the figure. *b*, At various times, samples of 25 μ l were precipitated with TCA and separated by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulphate⁹. The amount of G protein present was determined from scans of autoradiographs of the dried gels. The triangles and circles represent data from two experiments, run in identical conditions. The amounts of G protein shown are values obtained by normalisation to the amount of G protein made at the longest time point shown for each experiment, in order to permit the two experiments to be compared directly. Note that the concentration of ⁷mGp required to achieve satisfactory inhibition of initiation (> 99%) is very sensitive to the KCl concentration and that the time for chain completion can vary markedly ($\pm$ 50%) between independently prepared extracts of wheat germ.

protein in the G_1 form. Trypsinisation of vesicles after protein synthesis removes peptides from the extreme C-terminal region of G_1 , yielding a large fragment which migrates slightly faster than G_1 (but more slowly than G_0) on acrylamide gels (Fig. 2) and is protected from further tryptic degradation by the permeability barrier of the membrane vesicle^{8,9} wall through which G_1 is inserted. G_0 yields no analogous protected portion⁹. The extent to which the G protein was properly inserted into the membrane was, therefore, assayed as the fraction of radioactivity in the total G product ($G_0 + G_1$) preserved in the membrane-protected tryptic fragment which runs between G_0 and G_1 on gels. Even if all copies of G were inserted correctly, some loss of radioactivity would be expected since three C-terminal peptides containing ³⁵S-methionine radioactivity are removed by trypsin in the process of conversion of G_1 to the protected fragment⁹.

Figure 2 shows that when chains are completed in the presence of membranes after 6 min of synchronous synthesis in the absence of membranes, the resulting G protein is all in the G_1 , or glycosylated form, and yields a membrane-protected fragment on trypsinisation. A dramatic change takes place thereafter. Few of the G chains made for 9 min in the absence of membranes, and then completed in the presence of membranes, are glycosylated, and little if any membrane-protected fragment is observed (Figs 2, 3). Both glycosylation and proper insertion are lost coordinately and remarkably rapidly after 6 min. Indeed, both processes are lost within a span of 3 min, when the window in which initiation was permitted was 2 min. Consequently, prolonging synthesis of a G chain (which was initiated at 0 min) in the absence of membranes beyond the critical time of about 6 min by as little as 1 min will result in a marked reduction of proper insertion and of glycosylation. A fourfold increase in the concentration of membranes does not cause a significant increase in the critical time of about 6 min (not shown), indicating that this time is not related to the rate at which ribosomes can

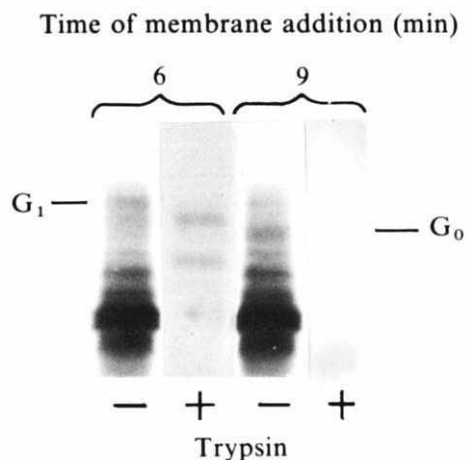


Fig. 2 Membrane addition experiment. At the indicated times, 50 μ l samples of the same protein synthesis incubation as Fig. 1 (Δ) were removed, 0.5 μ l of pancreatic microsomal membranes (prepared as before⁹; stock suspension had an A_{280} of 50, as measured in 1% sodium dodecyl sulphate) were added, and incubation was continued at 23 °C until 53 min had elapsed from the time mRNA was added. One half of each sample was precipitated with TCA⁹, while the other half was treated with trypsin-TPCK (0.5 mg ml⁻¹) at 23 °C for 15 min before TCA precipitation. Shown here is the autoradiogram of the gel of the translation products for the 6 and 9 min time points. Only the region of the gel containing G is presented; no time-dependent changes were observed for the other viral proteins. The gel was impregnated with scintillant³³ before drying, and the film was exposed for 40 d at -70 °C. The band below G_1 in the trypsin-treated lanes has methionine containing tryptic peptides of G (F. Katz, personal communication), and varies greatly in amount between experiments. Its significance is unclear.

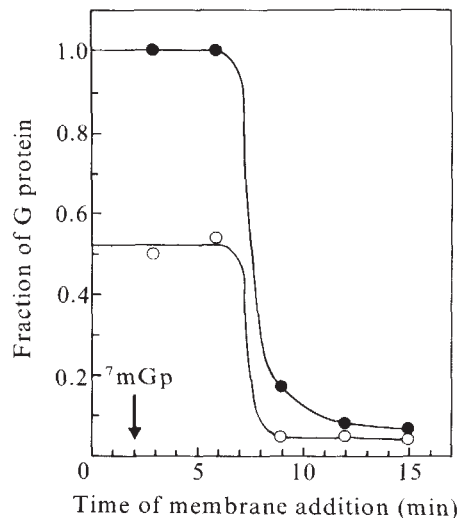


Fig. 3 Effect of the time of membrane addition on glycosylation and proper insertion. Same experiment as Fig. 2. The autoradiogram of the gel was scanned, and the amounts of G_0 and G_1 (lanes without trypsin treatment) and of the large membrane-protected tryptic fragment of G_1 (lanes with trypsin treatment) were determined from the peak heights. The extent of glycosylation ($\bullet$) is $G_1/(G_0 + G_1)$. The extent of protection of G by membranes ($\circ$) is the amount of fragment (running between G_0 and G_1) divided by $G_0 + G_1$, determined from the lane without trypsin of the same time point.

encounter membranes when the two are first mixed. This suggests that the critical time is dependent on some intrinsic property of the nascent G protein. This view is strongly supported by experiments in which the time for completion of G was about 23 min (using a different wheat germ preparation from Fig. 3), and the critical time for proper insertion was only 3 min. The critical time is therefore proportional to the time required for completion of the polypeptide.

If it is assumed that the growth of the G polypeptide is proportional to time, then the approximate length of the chain at the critical time can be determined. G has a molecular weight of ~ 63,000 (ref. 11) and therefore contains approximately 550 amino acid residues. In the experiment of Fig. 3, the chain was completed in 40 min, so about 80 residues would have been completed at the critical time. Of these, about 40 are present within the large subunit of the ribosome^{12,13}, so that only 40 residues would be available for any interaction with membrane vesicles. Synthesis of no more than 10–15 additional residues in the absence of membranes at this critical chain length results in the loss of proper insertion into the membrane and of glycosylation.

It can be concluded that both insertion into the membrane and glycosylation are closely coordinated with polypeptide chain synthesis, for membranes are required almost from the outset.

Sequence of glycosylation

The finding (Fig. 4) that protein synthesis in wheat germ extracts is unaffected by high concentrations of the detergent Triton X-100 has permitted the times of the glycosylation events to be determined in a simple experiment, and has led to the discovery of a novel intermediate in glycosylation. The lack of effect of Triton at concentrations as high as 1% (or possibly higher) is in marked contrast to the complete inhibition by the bile salt deoxycholate at much lower concentrations, as found for the reticulocyte system¹⁸. Concentrations of Triton as high as 1% would be expected to completely solubilise the pancreatic vesicle membranes, and might be expected to cause an abrupt inhibition of glycosylation by producing a random physical segregation of the nascent G chain (with attached ribosome), the glycosyltransferases, and the presumed lipid-linked sugar

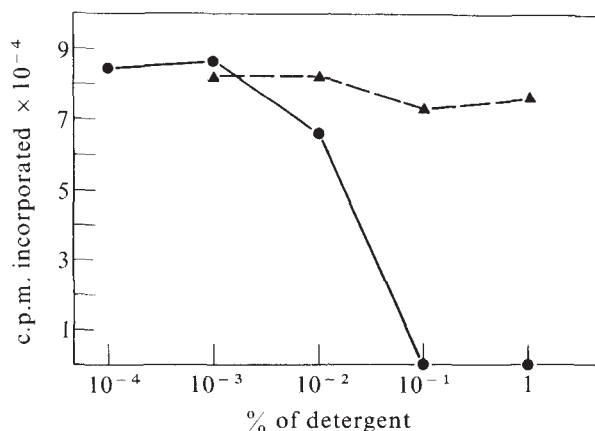


Fig. 4 Effect of detergents on protein synthesis. Conditions of protein synthesis were essentially the same as Fig. 1, except no ³mGp was added. Shown is the amount of TCA-insoluble protein made in 60 min per 2 μ l of reaction as a function of the concentration (w/v %) of either Triton X-100 (▲) or sodium deoxycholate (●).

substrates into different detergent micelles. Indeed, all of the G protein which results from translation of VSV mRNA in the presence of pancreatic membranes and 1% Triton X-100 is in the G₀ (unglycosylated) form (Fig. 5).

These results imply that the timing of glycosylation can be explored by synchronising protein synthesis with ³mGp as before but with the exception that membrane vesicles are present from the outset. At various times, samples are taken and Triton X-100 is added to inhibit further glycosylation, and the polypeptide chains are completed in the presence of the detergent. From the resulting distribution of G chains between glycosylated and unglycosylated forms as a function of the time of detergent addition, the times of glycosylation events can be inferred. Thus, when detergent is added before any glycosylation has taken place, no subsequent glycosylation can occur, so all of the resulting G will migrate in the G₀ position. When Triton is added after all glycosylation is completed, only the G₁ form will result.

Figure 5 shows that up to 9 min of synchronous synthesis, no glycosylation has taken place since only G₀ is observed. After 29 min, mainly G₁ is observed, so glycosylation is mainly complete well before the chain is finished. In between, and particularly around 20 min, a previously unknown form of G predominates, migrating on the gel in between G₀ and G₁. The amounts of these three forms as a function of the time of addition of detergent are shown in Fig. 6. The novel form of G, called G_{1/2}, is clearly a kinetic intermediate in the formation of G₁. This form binds specifically to columns of concanavalin A-Sepharose 4B, as judged by the methods used previously⁹ (data not shown), and is therefore (as expected from its intermediate mobility in gels) an intermediate in glycosylation.

The VSV virion glycoprotein contains two essentially identical oligosaccharide side chains¹⁴. As it is unlikely that both of the sites are glycosylated at the same time, and as the Asn-linked oligosaccharide would be derived by *en bloc* transfer from a lipid-linked oligosaccharide^{15,16}, an intermediate form in which one of the two sites is glycosylated would be expected. It was confirmed recently by analysis of the kinetics of degradation of G₁ by endoglycosidase H that G_{1/2} contains only one of the two oligosaccharides (J. Rothman, in preparation). This enzyme attacks each oligosaccharide at only one site, generating partial degradation products in proportion to the number of glycosylated Asn residues. It was found that G₁ was converted to G_{1/2}, which was then converted to G₀.

Since glycosylation is completed before the completion of the polypeptide backbone (Fig. 6), it is clear that the transfer of oligosaccharides is to the nascent chain. The rather precise kinetics of glycosylation probably indicate that each of the two

successive transfers of oligosaccharides to the nascent chain can only occur only after specific lengths of the nascent chain have been made because the timing of these events is determined only by the time required to complete G. When 23 min were required for completion of G chains (using a different wheat germ preparation from that used in Fig. 6), G_{1/2} was maximal at 12 min, and G₁ was half-maximal at 17 min, as compared with 22 min for G_{1/2} and about 30 min for G₁ when about 43 min were required for completion of G (Fig. 6).

Implications for the mechanism of insertion

Figure 7 presents a diagram of what is currently known of the sequence of events which occur during the synthesis of the VSV glycoprotein.

Our finding (Fig. 3) of a very early requirement for membranes implies that the insertion of G protein into the membrane has already begun when 80 or fewer residues are made, of which only 40 or fewer can be available for any interaction with membranes. The experiment only indicates the longest chain length which still permits subsequent insertion, so these figures place an upper limit on when insertion actually occurs. Glycosylation events are evidently not the basis for the early membrane requirement as both glycosylation steps occur well after the requirement for membranes is expressed (Fig. 7). Glycosylation is therefore not important for insertion of the membrane protein, at least in the initial stages. Presumably, insertion into the membrane proceeds continuously after it has begun.

In addition, our results strongly suggest that the nascent chain spans the lipid bilayer such that the protein is extruded across the membrane as it is being polymerised, as predicted from considerations of membrane asymmetry¹. This conclusion follows immediately from the fact that the chain is glycosylated while it is still attached (through tRNA) to the ribosome on the 'cytoplasmic' side of the membrane, providing it is assumed that the glycosylation steps take place on the opposite side of the membrane, within the vesicle.

Implications for the mechanisms of glycosylation

The belief that glycosylation is restricted to the extracytoplasmic surface of membranes is widespread and is based on the fact

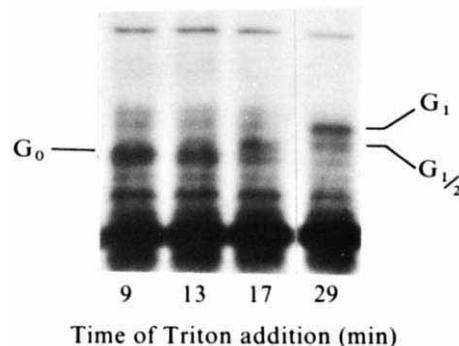


Fig. 5 Triton X-100 addition experiment. Protein synthesis was performed in the same conditions as Fig. 1, except the final volume was 0.95 ml and pancreatic microsomal membranes (20 μ l) were present before adding mRNA at time zero. At 2 min, 1 mM ³mGp was added. At each indicated time, a sample (25 μ l) was removed, 2.5 μ l of 10% (w/v) Triton X-100 was added, and incubation at 23 °C was continued until 53 min had elapsed from the time of addition of mRNA. Samples were then precipitated with TCA and electrophoresed as before⁹. Only the regions of the autoradiogram of the gel containing G are shown since no time-dependent changes in other viral proteins were observed. The gel was impregnated with scintillant³³ before drying, and was autoradiographed for 40 d at -70 °C. The band above G₁ was observed in the presence or absence of membranes, varied between experiments, and was not time-dependent. The G₁ chains required between 40 and 45 min for completion in the presence of membranes in this experiment, as determined by the methods of Fig. 1 (data not shown). The rate of elongation of proteins is therefore unaffected by the presence of membranes.

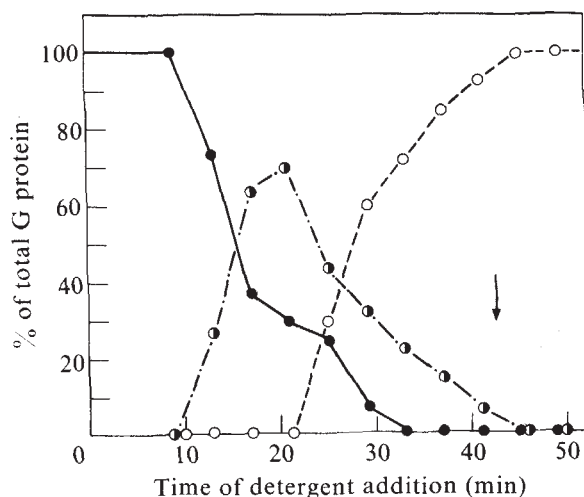
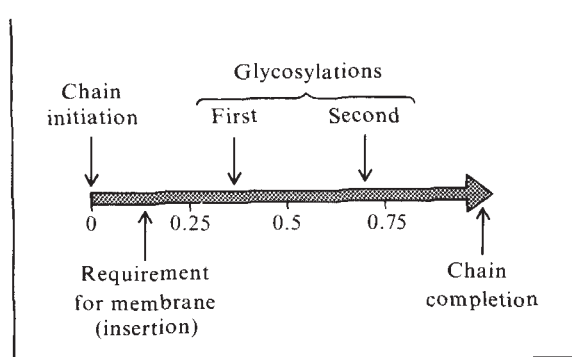


Fig. 6 Sequence of glycosylation. The amounts of G-specific translation products, G_0 , $G_{1/2}$ (the band between G_0 and G_1 in Fig. 5), and G_1 were determined for each time point from the peak heights of the corresponding bands in the acrylamide gel for the experiment which is partially presented in Fig. 5. The amount of each form is expressed as a % of the total of the three forms. G_0 (●), $G_{1/2}$ (●), G_1 (○). Arrow indicates polypeptide chain completion.

that membrane carbohydrate is found only on that side^{1,2}. Consistent with this, we have found (Fig. 3) that both glycosylation and proper insertion of the VSV glycoprotein are lost coordinately when membrane vesicles are added too late, in spite of the fact that the glycosylation steps themselves do not take place until well after this requirement for membranes is expressed (Fig. 7). This indicates that the late glycosylation depends on the early insertion of the nascent chain into the membrane on the 'cytoplasmic' side. This implies that the late glycosylation steps occur when the nascent chain of the G protein is present on the opposite, luminal side of the vesicle membrane.

We have presented here the first direct evidence that the nascent chain of a membrane glycoprotein is glycosylated. Kiely *et al.*¹⁷ showed that the majority of the carbohydrate is attached only to those peptide chains of ovalbumin (a secreted protein) that have been essentially completed. The timing of the two glycosylation steps of the VSV glycoprotein (Fig. 6) is remarkably precise. Each glycosylation takes place only after a specific length of the polypeptide chain is made. One explanation for this would be that the order and timing of the glycosylations is determined only by the order of the two acceptor residues in the amino acid sequence, that is, when the acceptor Asn residues first appear on the luminal side of the membrane.

Fig. 7 Sequence of events during the synthesis of G protein. The numbers below the arrow indicate the fraction of the chain completed, assuming the chain grows at a constant rate. The time of the first glycosylation is when G_0 is half of its original value, and the time for the second glycosylation is when G_1 is half maximal (Fig. 6).



Another plausible explanation would be that only when specific lengths of chain are made (and perhaps only for small intervals of chain length) is the nascent membrane protein in the appropriate conformation to act as an acceptor. In this case, an Asn would not generally be glycosylated at a time which corresponds to its position in the amino acid sequence. Such a conformation would presumably require that the Asn acceptor site of the nascent chain be in intimate proximity with the luminal surface of the lipid bilayer, in order to permit enzymatic transfer from the dolichol-oligosaccharide donor, located at the same interface.

Role of signal sequences

The considerations discussed above and elsewhere^{4,19} imply that the ribosomes which manufacture G protein are bound to the membrane by the nascent chain, if by no other interaction. Our results show for the first time that membrane-bound ribosomes engaged in the secretory process can be derived from free ribosomes after initiation of protein synthesis, a central aspect of the 'signal hypothesis'^{20,21}. Thus, when G chains are started in the absence of membranes on free ribosomes, and then subsequent initiation is blocked, these same ribosomes must become bound to the membrane in a functional configuration in order to permit the insertion and glycosylation of the resulting G protein which is observed when membranes are added at sufficiently early times (Fig. 3). Earlier work²¹ did not distinguish whether initiation occurred on added free ribosomes before or after they became bound to the membrane.

It then follows that these free ribosomes must possess a 'signal' which directs them to the correct membrane to bind in a functional way, permitting the nascent chain to be extruded through the membrane, as proposed in the 'signal hypothesis'^{20,21}. This signal could, in principle, be found in sequences of mRNA or of the nascent polypeptide itself. Support for the localisation of the 'signal sequence'²¹ to the N terminus of the nascent chain has come from the original finding of a precursor to the immunoglobulin light chain containing an additional sequence at the N terminus²⁰ and the subsequent finding that a wide spectrum of proteins secreted by pancreas possess essentially identical N-terminal extensions which are removed from the proteins in the endoplasmic reticulum²². Much additional work has shown that, quite generally, secreted proteins²³⁻³¹ (as well as a membrane ectoprotein³²) have similar N-terminal extensions some 15-30 residues in length. Unfortunately, direct evidence for the function of these extensions as the proposed signal sequences has been lacking.

Our findings now provide direct evidence for an N-terminal signal sequence which functions in the predicted manner²¹. Thus, the requirement for membranes during synthesis of G protein is not expressed until some 40 residues are found outside the ribosome, a length of chain just sufficient to contain any of the N-terminal extensions cited above. On the basis of the signal hypothesis, membranes could not be required until all of the signal sequence is translated and outside the large ribosomal subunit. The rapid loss of proper insertion of G when no more than 10-15 additional residues are made would then indicate that the signal sequence contained within the N-terminal 40 residues can only function during a short interval in the synthesis of the glycoprotein. There is as yet no evidence for a significant cleavage at the N terminus of G, but there is no obvious reason why an N-terminal signal sequence need be cleaved off.

The mechanisms used by the VSV glycoprotein to enter membranes can be expected to be relevant to most or all cellular ectoproteins since the limited coding capacity of the viral genome (five proteins, all structural components of the virion) requires that it utilises cellular factors for glycoprotein glycosylation and insertion. Indeed, the diverse sources of the components (plant ribosomes and animal membranes) which function together in the cell-free system indicate a high degree of evolutionary conservation of the mechanisms involved. The results obtained do not reflect unique properties of the particular

components used since the glycoprotein produced in the cell-free system is indistinguishable from the glycoprotein found in the endoplasmic reticulum of infected cells⁹. Further investigations using the cell-free system may provide more detailed insights into the mechanism by which the nascent chain crosses the membrane, and the relationship of signal sequences to this mechanism.

This research was supported by grants from the USNIH and the American Cancer Society. J. E. R. is a postdoctoral fellow of the Damon Runyon Foundation. We thank Dr. Richard Palmiter for useful discussions, Susan Froshauer and Martin Brock for technical assistance and Marianne Robotham for assistance with the manuscript.

Received 29 June; accepted 22 August 1977.

- Rothman, J. E. & Lenard, J. *Science* **195**, 743–753 (1977).
- Bretscher, M. S. *Science* **181**, 622–629 (1973).
- Lenard, J. & Compans, R. W. *Biochim. biophys. Acta* **344**, 51–94 (1974).
- Morrison, T. & Lodish, H. F. *J. biol. Chem.* **250**, 6955–6962 (1975).
- Grubman, M. J., Moyer, S. A., Banerjee, A. K. & Ehrenfeld, E. *Biochem. biophys. Res. Commun.* **62**, 531–538 (1975).
- Knipe, D. M., Baltimore, D. & Lodish, H. F. *J. Virol.* **21**, 1128–1139 (1977).
- Hunt, L. A. & Summers, D. F. *J. Virol.* **20**, 637–645 (1976); **20**, 646–657 (1976).
- Katz, F. N. & Lodish, H. F. (submitted).
- Katz, F. N., Rothman, J. E., Lingappa, V. R., Blobel, G. & Lodish, H. F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3278–3282 (1977).
- Hickey, E. D., Weber, L. A. & Baglioni, C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 19–23 (1976).
- Both, G. W., Moyer, S. A. & Banerjee, A. K. *Proc. natn. Acad. Sci. U.S.A.* **72**, 274–278 (1975).
- Blobel, G. & Sabatini, D. D. *J. Cell Biol.* **45**, 130–145 (1970).
- Malkin, L. I. & Rich, A. *J. molec. Biol.* **26**, 329–346 (1967).
- Robertson, J. S., Etchison, J. R. & Summers, D. F. *J. Virol.* **19**, 871–878 (1976).
- Waechter, C. J. & Lennarz, W. J. *A. Rev. Biochem.* **45**, 95–112 (1976).
- Pless, D. D. & Lennarz, W. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 134–138 (1977).
- Kiely, M. L., McKnight, G. S. & Schimke, R. T. *J. biol. Chem.* **251**, 5490–5495 (1976).
- Olsnes, S., Spaeren, U., Heiberg, R. & Pihl, A. *Biochem. J.* **130**, 297–299 (1972).
- Lodish, H. F. & Froshauer, S. *J. Cell Biol.* **74**, 358–364 (1977).
- Milstein, C., Brownlee, G. G., Harrison, T. M., & Mathews, M. B. *Nature new Biol.* **239**, 117–120 (1972).
- Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 835–851 (1975); **67**, 852–862 (1975).
- Devillers-Thierry, A., Kindt, T., Scheele, G. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 5016–5020 (1975).
- Burstein, Y. & Schechter, I. *Proc. natn. Acad. Sci. U.S.A.* **74**, 716–720 (1977).
- Shields, D. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2059–2063 (1977).
- Kemper, B., Habener, J. F., Ernst, M. D., Potts, J. T. & Rich, A. *Biochemistry* **15**, 15–19 (1976).
- Chan, S. J., Keim, P. & Steiner, D. F. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1964–1968 (1976).
- Strauss, A. W., Donohue, A. M., Bennett, C. D., Rodkey, J. A. & Alberts, A. W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1358–1362 (1977).
- Szczesna, E. & Boime, I. *Proc. Natn. Acad. Sci. U.S.A.* **73**, 1179–1183 (1976).
- Sussman, P. M., Tushinski, R. J. & Bancroft, F. C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 29–33 (1976).
- Dannies, P. S. & Tashjian, A. H., Jr *Biochem. biophys. Res. Commun.* **70**, 1180–1189 (1976).
- Inouye, H. & Beckwith, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1440–1444 (1977).
- Inouye, S., Wang, S. S., Sekizawa, J., Halegoua, S. & Inouye, M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1004–1008 (1977).
- Laskey, R. A. & Mills, A. D. *Eur. J. Biochem.* **56**, 335–341 (1975).

letters to nature

Radio haloes around BL Lacertae objects AO0235+164 and 4C03.59

THE rapid variations in intensity and linear polarisation of several BL Lacertae sources at centimetre and optical wavelengths have indicated the presence of multiple non-thermal components, often appreciably smaller than 1 pc (refs 1, 2). But there is no clear evidence for radio structures extended over kiloparsecs around any of the known variable BL Lac sources which have been scanned with narrow beams of a few arcs^{3–7}. The absence of extended radio structure in brighter and more variable BL Lac sources has also been inferred from their flat or inverted radio spectra and has led to the suggestion that these are extremely 'young' QSOs which have yet to develop a steady, non-thermal background^{8–10}. We report here the lunar occultation and interplanetary scintillation (IPS) observations of two BL Lac sources, AO0235+164 and 4C03.59 at 327 MHz. One of the two occultations of the former source, which is known to be an exceptionally violent variable at centimetre, infrared and optical wavelengths, took place in November 1975 near the epoch of its major outburst^{11–13}. These observations indicate that at 327 MHz, part of the emission of both these sources originates in components larger than 1 arc s.

The Arecibo occultation source AO0235+164 was identified with a 19-mag stellar object and was shown to be a radio and optical variable having an inverted radio spectrum and a featureless, steep optical spectrum^{3,16,17}. During late 1975, when this BL Lac source flared up dramatically, spectrometry revealed two optical absorption-line systems^{14,15} at redshifts of $z=0.524$ and 0.852 and also HI absorption¹⁵ at $z=0.524$. The strikingly rapid flux variations of this source impose extreme constraints on the incoherent synchrotron interpretation for rapidly varying extragalactic sources, if their redshifts are cosmological^{12,13,18}. The other BL Lac source 4C03.59 also has a steep optical spectrum^{3,19,20} but only a normal, straight radio spectrum²¹ with a slope of -0.6 and has been classified as a possible radio variable²². Visual images of both these BL Lac objects exhibit a jet-like nebulousity extending up to ~ 3 arc s from the stellar nucleus^{13,17,20}.

The lunar occultation and IPS observations at 327 MHz were made using the Ooty radio telescope. The methods of reducing such observations are described elsewhere^{23–25}. The results are given in Table 1, Figs 1 and 2. Occultation scans were observed for three position angles across AO0235+164 on 28 August and 18 November, 1975 and for four position angles across 4C03.59 between December 1973 and October 1974. The IPS observations were reduced using PKS1148–00 as calibrator, for which observations at Ooty have shown that its scintillation index rises to a maximum value of $85 \pm 5\%$ (S. Ananthakrishnan, personal communication).

After applying corrections for beam-broadening to the restored strip scans of Fig. 1, we estimate the half-power width of an equivalent gaussian source to AO0235+164 to be 3.0 ± 1.4 , 6.3 ± 1.8 and 3.7 ± 1.5 arc s in position angles of 21° , 66° and 140° , respectively. Thus AO0235+164 is resolved significantly at 327 MHz. We have also carried out IPS observations of this source on 10 days during April and May 1977, which indicate a scintillating component of size < 0.1 arc s containing a flux of 0.75 ± 0.15 Jy out of the total 1.07 ± 0.11 Jy at 327 MHz (Table 1). A simple interpretation of these measurements would be that at 327 MHz

Table 1 The observed parameters at 327 MHz

	AO0235+164	4C03.59
RA (1950)	02 ^h 35 ^m 52.61 $\pm$ 0.04s (02 35 52.63 $\pm$ 0.04) [†]	23 ^h 35 ^m 34.28 $\pm$ 0.03s (23 35 34.25 $\pm$ 0.03) [†]
Dec. (1950)	+16° 24' 04.5 $\pm$ 0.5'' (+16 24 04.3 $\pm$ 0.4) [†]	+03° 10' 11.6 $\pm$ 0.4'' (+03 10 11.8 $\pm$ 0.5) [†]
Flux density*	1.07 $\pm$ 0.11 Jy	4.3 $\pm$ 0.3 Jy
Flux of the IPS component	0.75 $\pm$ 0.15 Jy	0.4 $\pm$ 0.2 Jy
Size of the IPS component	< 0.1 arc s	0.25 $\pm$ 0.05 arc s

* The flux scale is as defined by Veron, Veron and Witzel²⁶.

[†] The values inside brackets are the centroid positions measured at 2.7 GHz using the RRE interferometer^{16,22} and are in excellent agreement with the positions of the identified optical objects, measured on the Sky Survey prints to an accuracy of ± 0.4 arc s in each coordinate^{16,27}.

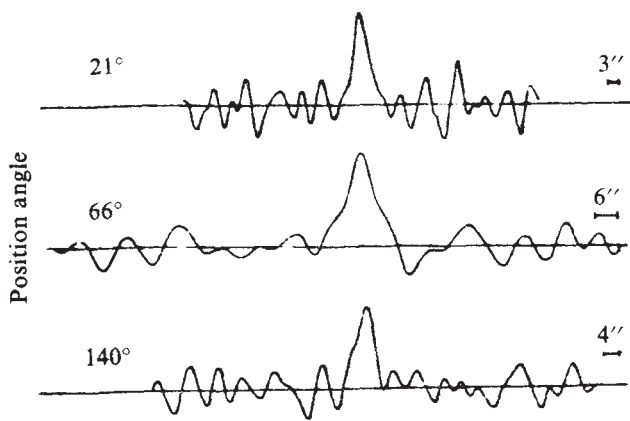
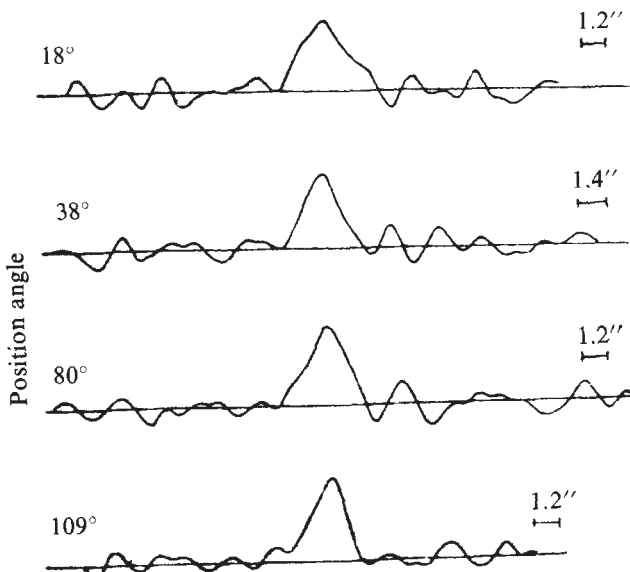


Fig. 1 The strip-brightness distributions across AO0235+164 are shown for the three indicated position angles. The resolutions are indicated by bars. Right ascension increases to the left. The scan in position angle 66° was observed on 18 November 1975 near the peak of the major radio outburst of this source^{11,12}.

the source consists of a core of size <0.1 arcs which accounts for nearly two-thirds of the total emission and an underlying extended component having a size of a few arc s. During VLBI observations²⁸, carried out in 1970 with a fringe spacing of 0.05 arc s, the source showed a maximum amplitude of 0.97 ± 0.27 Jy at 430 MHz, as compared with the total flux of 1.50 ± 0.16 Jy measured by Sutton *et al.* at 408 MHz²⁹. These two measurements indicate that either the entire flux is contained in a gaussian source of size ~ 0.02 arc s which was only partially resolved by the VLBI or, alternatively, that the measured visibility was due to an unresolved core of size <0.01 arc s and the remainder flux comes from a broad component. The latter interpretation would be consistent with the simple model suggested above.

After correcting the strip scans shown in Fig. 2 for beam-broadening, the half-power width of the source, 4C03.59(2335+031), is estimated to be 1.5 ± 0.5 arc s. The source profile in position angle 109° shows a faint but probably real extension towards east. This could be identified with the optical nebulosity of a few arc s which is seen extending from the stellar object identified with this source²⁰. In addition to the above structure, there is also

Fig. 2 Strip-brightness distributions across 4C03.59 are shown for the four indicated position angles. The resolutions are indicated by bars. Right ascension increases to the left.



evidence for a more compact, scintillating core contributing $10 \pm 5\%$ of the total flux at 327 MHz (Table 1).

The extended radio features of size ≥ 1 arc s detected in these two BL Lac sources have brightness temperatures exceeding $\sim 10^5$ K at 327 MHz. Hence their emission at metre wavelengths is likely to be predominantly non-thermal. Assuming that the absorption-line redshift $z=0.852$ for AO0235+164 represents its cosmological redshift and for Einstein-de Sitter model with $H=50$ km s⁻¹ Mpc⁻¹, the extended component of this source has a size of ~ 50 kpc and a minimum energy density of $\sim 10^{-10}$ erg cm⁻³ in the form of relativistic electrons and magnetic field.

Like many QSOs, the BL Lac source AO0235+164 exhibits multiple absorption systems¹⁴. The absorption could arise either in the intervening galaxies along the line of sight or within gas clouds located near its nucleus and being driven away due to its large radiation pressure^{30,31}. An important requirement in the latter model is that the gas clouds do not expand freely during their outflow, though the mechanism of their confinement is not clear³²⁻³⁴. The confinement may occur due to external pressure caused by the extended non-thermal component, provided the simple core-halo model of the source is valid. But, the present observations are also consistent with a more complex structure without a halo. It may also be noted that the mildly relativistic speeds of the outflowing absorbing clouds, as are inferred³⁵ from the multiple-redshift systems in QSOs and AO0235+164, are attainable only in a period of $\geq 10^5$ yr after the ejection³². It is interesting that a similar age of $\geq 10^5$ yr is also implied for AO0235+164 by the measured size of its extended component, taking this non-thermal component to be an expanded remnant of the past activity in its nucleus.

I thank Professor G. Swarup and Dr M. N. Joshi for their valuable comments.

GOPAL-KRISHNA

Radio Astronomy Centre,
Tata Institute of Fundamental Research,
P.O. Box 8, Ootacamund 643 001, India

Received 15 June; accepted 31 August 1977.

- Altschuler, D. R. & Wardle, J. F. C. *Mem. R. astr. Soc.* **82**, 1-67 (1976).
- Kinman, T. D. *Astrophys. J.* **205**, 1-5 (1976).
- Stein, W. A., O'Dell, S. L. & Strittmatter, P. A. *A. Rev. Astr. Astrophys.* **14**, 173-195 (1976).
- Kapahi, V. K. *Nature phys. Sci.* **234**, 49-50 (1971).
- Conway, R. G. & Stannard, D. *Mon. Not. R. astr. Soc.* **160**, 31-33 (1972).
- Colla, G. *et al. Astr. Astrophys.* **38**, 209-223 (1975).
- Argue, A. N., Kenworthy, C. M., Ryle, M. & Shakeshaft, J. R. *Nature Phys. Sci.* **241**, 139-139 (1973).
- Altschuler, D. R. & Wardle, J. F. C. *Nature* **255**, 306-310 (1975).
- Pollock, J. T. *Astrophys. J. Lett.* **198**, L53-L56 (1975).
- Usher, P. D. *Astrophys. J. Lett.* **198**, L57-L60 (1975).
- MacLeod, J. E., Aller, H. G. & Dent, W. A. *Nature* **260**, 751-752 (1976).
- Ledden, J. E., Aller, H. G. & Dent, W. A. *Nature* **260**, 752-754 (1976).
- Rieke, G. H., Grasdalen, G. L., Kinman, T. D., Hintzen, P., Wills, B. J. & Wills, D. *Nature* **260**, 754-759 (1976).
- Burbidge, E. M., Caldwell, R. D., Smith, H. E., Liebert, J. & Spinrad, H. *Astrophys. J. Lett.* **205**, L117-L120 (1976).
- Roberts, M. S. *et al. Astr. J.* **81**, 293-297 (1976).
- Argue, A. N., Kenworthy, C. M. & Stewart, P. M. *Astrophys. Lett.* **14**, 99-104 (1973).
- Spinrad, H. & Smith, H. E. *Astrophys. J.* **201**, 275-276 (1975).
- Terrell, J. *Astrophys. J. Lett.* **213**, L93-L97 (1977).
- Strittmatter, P. A., Carswell, R. F., Gilbert, G. & Burbidge, E. M. *Astrophys. J.* **190**, 509-514 (1974).
- Craine, E. R., Tapia, S. & Tarenghi, M. *Nature* **258**, 56-57 (1975).
- Wall, J. V. *Aust. J. Phys. Astrophys. Suppl.* no. 24 1-63 (1972).
- McEwan, N. J., Browne, I. W. A. & Crowther, J. H. *Mem. R. astr. Soc.* **80**, 1-59 (1975).
- Kapahi, V. K., Joshi, M. N., Subrahmanya, C. R. & Gopal-Krishna *Astr. J.* **78**, 673-680 (1973).
- Venkatakrishna, K. L. & Swarup, G. (in preparation).
- Bhandari, S. M., Ananthakrishnan, S. & Rao, A. P. *Aust. J. Phys.* **27**, 121-128 (1974).
- Veron, M. P., Veron, P. & Witzel, A. *Astr. Astrophys. Suppl.* **13**, 1-53 (1974).
- Singal, A. K., Gopal-Krishna & Venugopal, V. R. (in preparation (1977)).
- Galt, J. A., Broten, N. W., Legg, T. H., Leparskas, H. J. A. & Yen, J. L. *Mon. Not. R. astr. Soc.* **178**, 301-305 (1977).
- Sutton, J. M., Davies, I. M., Little, A. G. & Murdoch, H. S. *Aust. J. Phys. Astrophys. Suppl.* no. 33, 1-44 (1974).
- Mushotzky, R. F., Solomon, P. M. & Strittmatter, P. A. *Astrophys. J.* **174**, 7 (1972).
- Scargle, J. D. *Astrophys. J.* **179**, 705-719 (1973).
- Kippenhahn, R., Perry, J. J. & Röser, H.-J. *Astr. Astrophys.* **34**, 211-224 (1974).
- Mathews, W. G. *Astrophys. J.* **189**, 23-31 (1974).
- Weymann, R. *Astrophys. J.* **208**, 286-291 (1976).
- Rees, M. J. *Astrophys. J. Lett.* **160**, L29-L32 (1970).

A constraint on the universal baryon density from the abundance of ${}^7\text{Li}$

THE observed interstellar abundance of ${}^2\text{H}$ has been used^{1,2} to estimate the mean baryon density (ρ_b) of the Universe. This follows, because (1) there is no plausible source for ${}^2\text{H}$ other than the primordial big bang and (2) the production of ${}^2\text{H}$ in a standard big bang decreases rapidly with increasing ρ_b . If one then assumes that all ${}^2\text{H}$ was formed in a big bang, the observed abundance² of this nuclide requires a value of ρ_b sufficiently low¹ that, for a cosmological constant $\Lambda=0$, the present expansion of the Universe will continue forever and the Universe is open. A major weakness in this argument is that another source of ${}^2\text{H}$ may be found. It has been suggested, for example, that ${}^2\text{H}$ could be made in shock waves accompanying a supernova explosion; this now seems unlikely³, but other mechanisms will certainly be suggested, so that it is important to obtain confirmation of the above conclusion. The predicted production of ${}^7\text{Li}$ in a big bang² varies rapidly with ρ_b and could be used to estimate ρ_b if the fraction of the observed ${}^7\text{Li}$ made in the big bang were known. Unfortunately there are many possible sources⁴ of ${}^7\text{Li}$ and such estimates must be regarded with scepticism. In this note we point out that ${}^7\text{Li}$ can be used to place an upper limit on ρ_b , even if other production mechanisms are important, and that this limit also strongly favours an open universe. This possibility arises because the big bang production of ${}^7\text{Li}$ increases with increasing ρ_b (for

$\rho_b > 10^{-31}$) so that an upper limit is obtained by attributing all of the observed ${}^7\text{Li}$ to the big bang.

We have adopted here Boesgaard's⁵ value of the Li abundance which yields⁶ a fractional abundance by mass of ${}^7\text{Li}$, $X_7 = 5 \times 10^{-9}$. Assuming the big bang must not synthesise more than this amount leads to $\rho_b \leq 1.1 \times 10^{-30} \text{ g cm}^{-3}$. As is shown in Fig. 1, this is substantially less than the critical value ρ_c necessary to close a $\Lambda=0$ Friedman universe.

The uncertainty in X_7 is perhaps a factor of two; the meteoritic value⁶, for example, is $X_7 = 8 \times 10^{-9}$. Substantially larger values have been seen⁵ in a small number of red giant stars, but these values presumably reflect a local production mechanism. Allowing for a factor of two uncertainty gives an upper limit closer to ρ_c , but still favouring an open and forever expanding universe.

The existence of mechanisms which destroy ${}^7\text{Li}$ weakens the limit on ρ_b since the big bang may then have made more ${}^7\text{Li}$ than is now observed; conversely, discovery of additional sources of ${}^7\text{Li}$ strengthens the limit. Astration of primordial material is presumably the most important destruction process. Estimates of the fraction of matter which has passed through stars are rather uncertain but are typically about 0.5. It has been pointed out^{4,7}, however, that infall of primordial material from the galactic halo may be significant and would tend to compensate for the effects of astration for those nuclei produced in the big bang. Other sources of ${}^7\text{Li}$ are generally rather speculative⁴, except for production in the cosmic rays which yields roughly 10% of the observed ${}^7\text{Li}$. Since these various effects tend to offset each other, the observed value of X_7 seems reasonable but subject to uncertainty.

If it is a good approximation to ignore both astration and sources of ${}^2\text{H}$ and ${}^7\text{Li}$ other than the big bang, their observed abundances each separately determine the density. An estimate based on the ${}^2\text{H}$ abundance X_D is shown in Fig. 1, and is in good agreement with the density obtained from ${}^7\text{Li}$. Effects of astration would tend to worsen this agreement. Thus when other possible contributions to ${}^7\text{Li}$ are better understood, the requirement that the big-bang contribution to X_7 and X_D yield the same value of ρ_b may be a strong constraint on allowable astration.

We assumed above that the cosmological constant $\Lambda=0$. While this is consistent with the available data, a non-zero value cannot be excluded, except on aesthetic grounds, and its effects must be considered. It has been found¹¹ that for reasonable values of Λ , the limits on ρ_b from the ${}^2\text{H}$ and ${}^7\text{Li}$ abundances are essentially unchanged. But, the simplest relationship between ρ_b and the curvature and evolution of the Universe is no longer valid¹¹.

In summary, the simplest and most straightforward assumptions concerning the origin of ${}^7\text{Li}$ and the nature of the big bang expansion require an upper limit for the present universal density of $\rho_b = (1.1 [+1.1 \text{ or } -0.4]) \times 10^{-30} \text{ g cm}^{-3}$. Given that the Universe is indeed a Friedman universe with zero cosmological constant, the agreement between the present limit and that based on ${}^2\text{H}$ strongly supports the conclusion of Gott *et al.*¹ that the Universe is open and will continue to expand forever.

This research was supported by the USNSF.

Note added in proof: It has come to our attention that conclusions similar to those reached here have been discussed by G. Steigman at the Harvard Neighborhood Meeting on Cosmology, October 1975.

SAM M. AUSTIN
C. H. KING

Cyclotron Laboratory and Physics Department,
Michigan State University,
East Lansing, Michigan 48824

Received 27 June; accepted 15 August 1977.

- Gott, J. R., Gunn, J. E., Schramm, D. N. & Tinsley, B. M. *Astrophys. J.* **194**, 543–553 (1974).
- Schramm, D. N. & Wagoner, R. V. *A. Rev. nucl. Sci.* (in the press).
- Epstein, R., Arnett, W. D. & Schramm, D. N. *Astrophys. J. Suppl.* **31**, 111–141 (1976).
- Reeves, H. *A. Rev. Astr. Astrophys.* **12**, 437–469 (1974).
- Boesgaard, A. M. *Pub. astr. Soc. Pac.* **88**, 353–66 (1976).
- Cameron, A. G. W. *Space Sci. Rev.* **15**, 121–46 (1973).
- Audouze, J. & Tinsley, B. M. *Astrophys. J.* **192**, 487–500 (1974).
- Wagoner, R. V. *Astrophys. J.* **179**, 343–360 (1973).
- Woody, D. P., Mather, J. C., Nishioka, N. S. & Richards, P. L. *Phys. Rev. Lett.* **34**, 1036–1039 (1975).
- York, D. G. & Rogerson, J. B. Jr *Astrophys. J.* **203**, 378–385 (1976).
- Tinsley, B. M. *Phys. Today* **30**, 32–38 (1977).

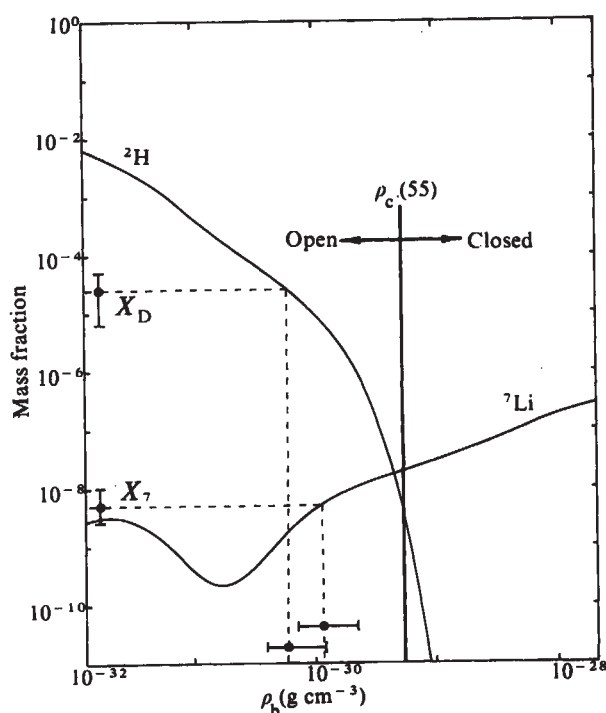


Fig. 1 Abundances of ${}^2\text{H}$ and ${}^7\text{Li}$ produced in a standard big bang (adapted from ref. 8. The present black body temperature is taken to be 2.90 K, see ref. 9.) The vertical line labelled ρ_c (55) is the density necessary to close a Friedman universe with $\Lambda=0$, if $H_0=55 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (in general $\rho_c=5.7 \times 10^{-30} (H_0/55)^2$). The point labelled X_7 is the mass fraction of ${}^7\text{Li}$ corresponding to the abundance given by Boesgaard⁵, while that labelled X_D is the mass fraction of ${}^2\text{H}$ from the summary of ref. 2. (This latter value is smaller than that used by Gott *et al.*¹, mostly because they include an estimate of the effects of astration). The uncertainty indicated for X_7 is a factor of two in either direction while that for X_D covers the range from a factor of four smaller to a factor of two larger¹⁰. Corresponding values of ρ_b and their uncertainties are also shown. The value of ρ_b determined from the ${}^7\text{Li}$ abundance is only an upper limit if there are significant sources of ${}^7\text{Li}$ other than the big bang.

Quantitative analysis of the Dermott–Gold theory for Uranus's rings

DERMOTT and Gold¹ have attempted to explain the locations of Uranus's rings in terms of resonances between ring particles and pairs of satellites, such that each particle librates about consecutive conjunctions of the satellites. The libration argument is given by

$$K = \lambda_1 - (p+1)\lambda_2 + p\lambda_3$$

where $\lambda_i = n_i t + \text{constant}$, $i = 1, 2, 3$, are the mean longitudes of the ring particle and the satellites, and p is an integer. Despite the apparent success of the theory in predicting the main observed features of the rings, it is necessary to investigate the strength of the resonances in more detail before any conclusions can be drawn. This letter summarises the outcome of such an investigation which supplements the previous largely qualitative analysis¹.

Aided by a study by Wilkens² of possible three-body resonances involving one minor and two major planets, I derived an equation of motion analogous to that of a pendulum

$$\frac{d^2 K}{dt^2} = c \sin K$$

where c is found to be roughly proportional to $m_2 m_3 (a_2/a_3)^p$, that is the product of the masses of the two satellites times the ratio of their semi-major axes to the p -th power. In Table 1, the first column gives the mass products adopted by Dermott and Gold for various pairs of satellites. The second column contains the range of values for $p = (n_1 - n_2)/(n_2 - n_3)$ required to span the observed width of the ring system—roughly 45,000–52,000 km. The corresponding range of c -values and of libration periods ($P = 2\pi \cdot c^{-1}$) are given in the final two columns (the quoted periods are for zero-amplitude librations; the period increases slowly with the amplitude, being twice as large for 164° as for 0°). Since $c > 0$, librations can occur only about $K = 180^\circ$.

On the basis of the magnitudes of the product $m_2 m_3$ and the resulting ring pattern, Dermott and Gold concluded that the observed pattern is probably due primarily to the Ariel–Titania and Ariel–Oberon pairs. From the values of c , however, we now see that the correction factor $(a_2/a_3)^p$ makes Miranda play the key role rather than Ariel, in spite of the small mass of the former. Furthermore, the single combination Miranda–Ariel dominates over all the others in terms of efficiency and predicts ring radii for $p = 7, 8$ and 9 which are very close to the radii obtained by Marsden³ for the ϵ_2 , γ , and α rings, respectively. Therefore, when applied to this and other satellite pairs, the Dermott–Gold mechanism, as modified here, seems to have the capability of explaining the observed ring locations. It is not clear, though, how much statistical weight this argument should be given, one problem being that the model may predict too many rings—which, incidentally, provides hope for a future observational test. The relatively low p values of the pairs involving Miranda may also add some new credibility to the theory. It is interesting to speculate that even lower values for p —and higher ones for c —would be attainable for a hypothetical pair of satellites involving Miranda and an as yet undiscovered satellite inside it. After all, the 80,000-km wide gap between the rings and Miranda is not likely to be entirely empty.

On the other hand, the very long libration periods derived in Table 1, when compared with revolution periods of only about a third of a day for the ring particles, put very tight constraints on the semi-major axes of librating particles. A radial displacement of the order of $\Delta a_1 = a_1 \cdot \sqrt{c/n_1} \leq 50$ m is enough to disrupt the extremely weak particle–satellite–satellite coupling

—and there are several competing forces that might set up a disturbance of that magnitude. Elliot *et al.*⁴ have estimated ring widths of from 12 to 85 km and particle sizes < 6 km. The theory cannot account for these observations unless the particles are assumed to have slightly eccentric orbits. One important question not yet answered is whether the restoring force of the particle 'pendulum' is sufficiently strong to overcome the Poynting–Robertson effect for reasonable particle sizes.

Table 1 Ranges of p , c , and $P = 2\pi c^{-1}$

Satellites	$m_2 m_3 \times 10^{11}$	p	$c(1,000 \text{ yr})^{-2}$	$P(1,000 \text{ yr})$
Miranda, Ariel	19	7–9	53–18	0.87–1.5
Miranda, Umbriel	5	4–6	5.0–1.1	2.8–5.9
Miranda, Titania	32	3–5	6.9–0.69	2.4–7.5
Miranda, Oberon	25	3–4	1.8–0.32	4.7–11
Ariel, Umbriel	99	15–20	4.6–0.73	2.9–7.4
Ariel, Titania	580	8–11	2.1–0.15	4.4–16
Ariel, Oberon	447	7–10	0.32–0.01	11–62
Umbriel, Titania	170	20–26	$(88-3.7)10^{-4}$	67–320
Umbriel, Oberon	131	15–19	$(63-2.1)10^{-5}$	$(2.5-14)10^2$
Titania, Oberon	770	66–83	$(56-0.32)10^{-8}$	$(0.84-11)10^4$

In conclusion, although sceptical of the Dermott–Gold theory, despite its ingenuity and elegance, I believe that a decisive test of the theory has to await further observational details of Uranus's rings. In the meantime, it would seem advisable to look for alternative ring theories. For example, Colombo (personal communication) points out that it may be of significance that the period associated with the ϵ ring is in the approximate ratios 5:4, 6:5, 9:8, and 11:10 to those associated with the α , β , γ , and δ rings. One should not discount the possibility, either, that the rings may be transient features not locked into any kind of resonance at all.

This work was supported by NASA grant NGR 09-015-213.

KAARE AKSNES

Center for Astrophysics,
Harvard and Smithsonian Observatories,
Cambridge, Massachusetts 02138

Received 5 August; accepted 31 August 1977.

1. Dermott, S. F. & Gold, T. *Nature* **267**, 590–593 (1977).
2. Wilkens, A., & Sitzungsber., D., *Math-Natn. Abt. Bayer. D., Akad. Wiss. D., Zu Munchen*, **1**, 71–101 (1933).
3. Marsden, B. G. *IAU Circ. No. 3061* (1977).
4. Elliot, J. L., Dunham, E., & Mink, D. *Nature* **267**, 328–330 (1977).

Revenge of tiny Miranda

DERMOTT and Gold¹ have proposed a resonance model for the rings of Uranus. They assume the rings are, in fact, arcs composed of small particles librating about stable resonances determined by pairs of satellites, either Ariel and Titania or Ariel and Oberon. Dermott and Gold dismiss as insignificant resonances involving tiny Miranda. We report here that, by a wide margin, the strongest resonances are all associated with Miranda. Furthermore, we show that the hypothesis that the rings are made up of librating particles, while original and ingenious, is incorrect.

Before considering the quantitative analysis of the resonance model, we make two minor points. First, when allowance is made for the orbital motion (assumed prograde) of the ring material between occultations, it is found that the two occulting regions of each ring were physically only 35° (α ring) to 46° (ϵ ring) apart. Thus, only short arcs are required to fit the observations. Second, while large systematic errors may remain in the calculated absolute ring radii, the spacings between the rings are well determined and must be accurately predicted by a resonance theory. In Table 1, the spacings predicted by Dermott and Gold are compared with those deduced by Elliot *et al.*² and Marsden³. The discrepancies seem to be larger than observational error and exhibit no systematic trend.

We have calculated the strengths of the resonances between a ring particle in circular orbit and either a single satellite or a pair of

satellites. The resonance strength is expressed as the magnitude of the resonant term in the disturbing potential, R , felt by the ring particle. Figure 1 shows the strengths of the most important resonances in the radial range $42 \times 10^3 \text{ km} < a < 54 \times 10^3 \text{ km}$, which spans the ring radii. The only two-body resonances included in Fig. 1 are the 4:1 and 5:1 resonances with Miranda. Two-body resonances involving other satellites also lie in the range of the rings. The strengths, however, are too small for them to show in Fig. 1.

The strongest resonances are the two-body resonances with Miranda which occur where

$$4n_M - n - 3 \frac{d\tilde{\omega}_M}{dt} = 0$$

or

$$4n_M - n - \frac{d\tilde{\omega}_M}{dt} - \frac{d\Omega}{dt} - \frac{d\Omega_M}{dt} = 0$$

Here, n_M and n are the mean motions of Miranda and the ring particle, Ω_M and Ω are the longitudes of their ascending nodes, and $\tilde{\omega}_M$ is the longitude of Miranda's periapse. The resonance strengths are given by Brouwer and Clemence⁵.

$$R_1^4 = \frac{GM}{48a} \frac{m_M}{M} \frac{a}{a_M} e_M^3 \left[-256\alpha + \left(142 + 114\alpha \frac{d}{d\alpha} + 21\alpha^2 \frac{d^2}{d\alpha^2} + \alpha^3 \frac{d^3}{d\alpha^3} \right) b_{1/2}^{(1)}(\alpha) \right]$$

and

$$R_2^4 = \frac{GM}{4a} \frac{m_M}{M} \frac{a}{a_M} e_M (\sin i/2)^2 \left[\left(8 + \alpha \frac{d}{d\alpha} \right) b_{3/2}^{(2)}(\alpha) \right]$$

where M is the mass of Uranus, m_M and e_M the mass and eccentricity of Miranda, i the mutual inclination of the ring and Miranda's orbit, and $\alpha \equiv a/a_M$.

The functions $b_s^{(i)}$ are Laplace coefficients.

There are three 5:1 resonances, corresponding to

$$5n_M - n - 4 \frac{d\tilde{\omega}_M}{dt} = 0$$

$$5n_M - n - 2 \frac{d\tilde{\omega}_M}{dt} - \frac{d\Omega}{dt} - \frac{d\Omega_M}{dt} = 0$$

and

$$5n_M - n - 2 \frac{d\Omega}{dt} - 2 \frac{d\Omega_M}{dt} = 0$$

Using formulae given by Peirce⁶, we obtain the resonance strengths:

$$R_1^5 = \frac{GM}{384a} \frac{m_M}{M} \frac{a}{a_M} e_M^4 \left[-3125\alpha + \left(1569 + 1556\alpha \frac{d}{d\alpha} + 402\alpha^2 \frac{d^2}{d\alpha^2} + 36\alpha^3 \frac{d^3}{d\alpha^3} + \alpha^4 \frac{d^4}{d\alpha^4} \right) b_{1/2}^{(1)}(\alpha) \right]$$

$$R_2^5 = \frac{GM}{16a} \frac{m_M}{M} \frac{a}{a_M} e_M^2 (\sin i/2)^2 \left[\left(85\alpha + 20\alpha^2 \frac{d}{d\alpha} + \alpha^3 \frac{d^2}{d\alpha^2} \right) b_{3/2}^{(2)}(\alpha) \right]$$

$$R_3^5 = \frac{3GM}{8a} \frac{m_M}{M} \frac{a}{a_M} (\sin i/2)^4 \left[\alpha^2 b_{5/2}^{(3)}(\alpha) \right]$$

where a and α now refer to the position of the 5:1 resonance. All of these strengths depend sensitively on the uncertain values of e_M and i . Greenberg⁴ has determined that $e_M \simeq 0.012$ and $i \simeq 4^\circ$, though

he regards the latter as "extremely model dependent". Using these values, and noting, therefore, that only R_1^4 and R_1^5 are likely to be reasonably accurate, we obtain

$$R_1^4 \simeq 6.6 \times 10^{-13} \frac{GM}{a}$$

$$R_2^4 \simeq 1.1 \times 10^{-11} \frac{GM}{a}$$

$$R_1^5 \simeq 9.1 \times 10^{-15} \frac{GM}{a}$$

$$R_2^5 \simeq 3.0 \times 10^{-13} \frac{GM}{a}$$

and

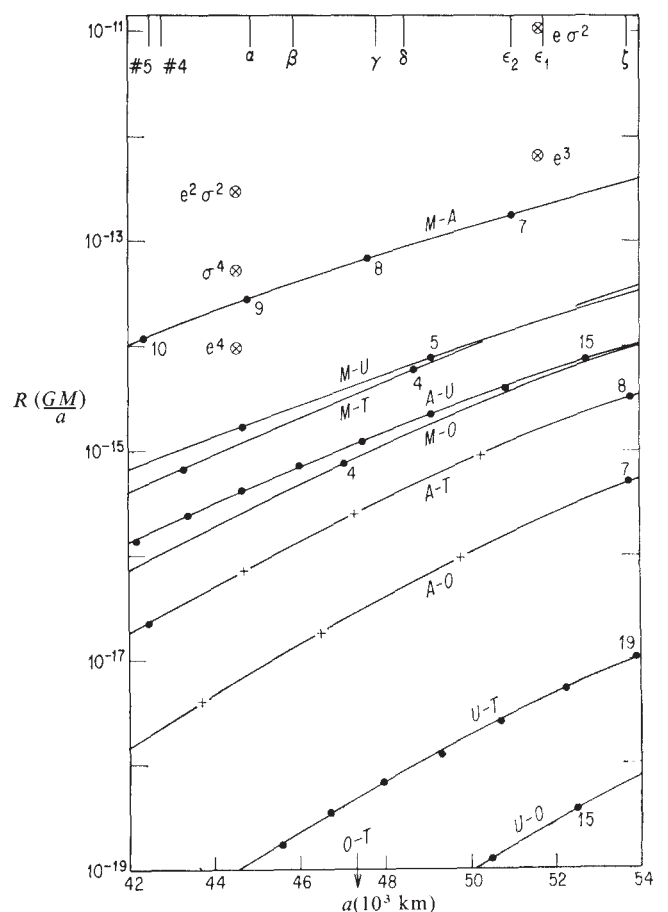
$$R_3^5 \simeq 5.2 \times 10^{-14} \frac{GM}{a}$$

Next, we consider three-body resonances of the form

$$qn - (p+q)n_B + pn_A = 0$$

where p and q are integers and n the mean motions with the subscripts A and B denoting the outer and inner satellites respectively. Both satellite orbits are assumed to be circles and to be outside the rings. The dominant resonant term in the disturbing potential arises as follows. Outer satellite A perturbs the orbit of inner satellite B, producing oscillatory variations in its radius and longitude with frequencies $p(n_A - n_B)$. The potential at the ring due

Fig. 1 The strengths of two- and three-body resonances in the neighbourhood of Uranus's rings. $\otimes$, The 4:1 and 5:1 resonances with Miranda discussed in the text and identified by $e = e_M$ and $\sigma = \sin(i/2)$; $\bullet$, the three-body resonances for $q=1$ and values of p as indicated. The resonances previously¹ associated with the rings are shown as +.



to B moving on its perturbed orbit contains terms with frequencies $qn - (p + q)n_B + pn_A$. These include the resonant term and many short period terms. Additional, but smaller (by a factor $\lesssim 0.1$), resonant terms arise from the attraction of A moving on an orbit perturbed by B. Finally, there are much smaller resonant terms due to the interaction between the direct perturbations of the ring by each satellite moving on its unperturbed circular orbit. The dominant resonant term is easily shown to be

$$\mathbf{R}_{pq} = \frac{GM}{2a} \frac{m_A m_B}{M^2} \frac{a_A a^2}{a_B^3} \frac{n_A^2}{[p^2(n_B - n_A)^2 - n_B^2]} \times \\ \times \left\{ h_q(\alpha) \left[f_p(\beta) - \frac{2n_B}{p(n_B - n_A)} g_p(\beta) \right] + g_q(\alpha) \times \right. \\ \left. \times \left[\left(1 + \frac{3n_B^2}{p^2(n_B - n_A)^2} \right) g_p(\beta) - \frac{2n_B}{p(n_B - n_A)} f_p(\beta) \right] \right\} \quad (9)$$

where a is orbital radius, m the satellite mass and $\alpha \equiv a/a_B$, $\beta \equiv a_B/a_A$. The functions f_k , g_k and h_k may be expressed in terms of Laplace coefficients (ref. 5) $b_s^{(j)}$ by

$$f_k(x) = x b_{3/2}^{(k)}(x) - \frac{1}{2} [b_{3/2}^{(k+1)}(x) + b_{3/2}^{(k-1)}(x)] + \delta_{k1} \quad (10)$$

$$g_k(x) = \frac{1}{2} [b_{3/2}^{(k-1)}(x) - b_{3/2}^{(k+1)}(x)] - \delta_{k1} \quad (11)$$

$$h_k(x) = \frac{1}{2} [b_{3/2}^{(k-1)}(x) + b_{3/2}^{(k+1)}(x)] - b_{3/2}^{(k)}(x)/x + 2\delta_{k1} \quad (12)$$

We have calculated the strengths of all $q = 1$ resonances due to each pair of satellites using power series approximations for the Laplace coefficients and satellite parameters from Greenberg⁴. The results are displayed in Fig. 1, with a smooth curve connecting the resonances due to each satellite pair. It is apparent that resonances involving tiny Miranda, the innermost satellite, dominate the field. The Ariel–Titania and Ariel–Oberon resonances advocated by Dermott and Gold¹ are relatively weak and presumably incapable of determining the ring locations.

Figure 1 shows that the strong Miranda–Ariel resonances lie close to four of the rings: Millis *et al.*'s ring 5 (ref. 7) and Elliot *et al.*'s rings α , γ and ϵ_2 , (ref. 2). Differences between Marsden's³

Table 1 Observed compared with resonant ring spacings

Ring pair	Spacing (km)*		Resonance model ¹	Difference (km)
	Marsden ³	Elliot <i>et al.</i> ²		
α, β	949	926	1,026	+89
γ, δ	680	672	800	+124
ϵ_1, ϵ_2	632	~ 600	483	~ -130

*Calculated values are averages of occultation entry and exit results.

calculated positions for these rings and the resonances corresponding to $p = 10, 9, 8$ and 7 are 155, 96, 157 and 111 km, respectively, for an assumed $J_2 = 0.013$ for Uranus. As a consequence of the approximate Laplace relation satisfied by the mean motions of Miranda, Ariel and Umbriel⁸, several weaker Miranda–Umbriel and Ariel–Umbriel resonances also approximately coincide with the above rings. Furthermore, the 4:1 resonance with Miranda is located about 114 km inside Marsden's³ position for the ϵ_1 ring.

Despite the near coincidences between the locations of the strongest resonances and some of the rings, the rings cannot be made up of librating material. The maximum radial width of an arc of librating particles is

$$W = 8 \left(\frac{aR}{3GM} \right)^{1/2} a \quad (13)$$

From the resonance strengths, we obtain $W \approx 0.7$ km for the 4:1 resonance and $W \lesssim 0.07$ km for the strongest Miranda–Ariel resonances. By comparison, the widths of the observed rings range from 1–10 km for the inner rings and 30–100 km for the ϵ ring or

rings. Furthermore, libration occurs about a relative maximum of the potential energy in the frame rotating with the resonant mean motion. Thus, inelastic collisions among the particles would be destabilising and lead to the dissolution of a compact arc of librating material. We conclude that if the ring positions are determined by resonances, the control is more subtle than previously suggested. One possibility is that the rings are the crests of nonlinear density waves in an optically thin disk of particles. Calculations of the resonant excitation of density waves in Saturn's rings have shown that even weak resonances produce nonlinear waves⁹.

We thank G. Colombo for helpful discussions and acknowledge the support of NASA Grant NGL 05-002-003.

PETER GOLDBREICH
PHILIP NICHOLSON

Division of Geological and Planetary Sciences,
California Institute of Technology,
Pasadena, California 91125

Received 1 August; accepted 31 August 1977.

1. Dermott, S. F. & Gold, T. *Nature* **267**, 590–593 (1977).
2. Elliot, J. L., Dunham, E. & Mink, D. *Nature* **267**, 328–330 (1977).
3. Marsden, B. G. *IAU Circ.* No. 3061 (1977).
4. Greenberg, R. *Icarus* **24**, 325–332 (1975).
5. Brouwer, D. & Clemence, C. M. *Methods of Celestial Mechanics* 490–494 (Academic, New York, 1961).
6. Peirce, B. *Astr. J.* **1**, 1–8, (1849).
7. Millis, R. L., Wasserman, L. H. & Birch, P. V. *Nature* **267**, 330–331 (1977).
8. Sinclair, A. T. *Mon. Not. R. astr. Soc.* **171**, 59–72 (1975).
9. Goldreich, P. & Tremaine, S. *Icarus* (submitted).

Identification of water frost on Callisto

THE 3.1- μm absorption band of water frost has been tentatively identified in spectra of the Galilean satellites Ganymede and Callisto. This is the first positive evidence for water frost on the surface of Callisto. The depth of this absorption band on both satellites is about twice as deep as is inferred from the depths of the water frost bands at 1.5 and 2.0 μm . Previous high resolution spectral studies ($\Delta\lambda/\lambda \sim 0.02$) of the Galilean satellites in the 1.0–2.5 μm region have identified water frost on the surfaces of Europa (J II) and Ganymede (J III) from the presence of characteristic water frost absorption bands at 1.5 and 2.0 μm ^{1–3}. All studies failed to positively identify water frost on either Io (J I) or Callisto (J IV) but were used to set an upper limit of 5–25% (surficial coverage) for the amount of water frost on Callisto¹. The spectrum of water frost also has a characteristic band at about 3.1 μm which is stronger than the ones at shorter wavelengths^{3–5}. Previous studies have shown a decreased reflectance between 3 and 4 μm for Callisto, but water frost could not be positively identified because of low spectral resolution^{6,7} and poor signal-to-noise and the problem of using the Moon as a standard because of the onset of lunar thermal emission beyond 2.5 μm (refs 1,2).

Broadband (J, K, and L) and narrowband (3.0–3.8 μm) observations of Ganymede and Callisto were made on the nights of 7 November 1976 and 29 November 1976 (UT) with the Mount Lemmon 28-inch infrared telescope which

Table 1 Flux calibration

Wavelength (μm)	$\Delta\lambda$ (μm)	Zero magnitude flux density	
		($\text{W cm}^{-2} \mu\text{m}^{-1}$)	($\text{W m}^{-2} \text{Hz}^{-1}$)
1.25 (J)	0.3	3.40×10^{-13}	1.77×10^{-23}
2.22 (K)	0.5	4.14×10^{-14}	6.80×10^{-24}
3.03	0.07	1.28×10^{-14}	3.92×10^{-24}
3.12	0.1	1.10×10^{-14}	3.57×10^{-24}
3.43	0.2	7.62×10^{-15}	2.99×10^{-24}
3.45 (L)	0.9	7.43×10^{-15}	2.58×10^{-24}
3.73	0.5	5.56×10^{-15}	2.46×10^{-24}

J and K values from Low and Rieke⁸. All others are interpolated values from the K and L values of Low and Rieke⁸.

employs a rocking secondary mirror to chop the beam 25 arcs on the sky. The infrared observational techniques and equipment are described elsewhere⁸. The filters used are listed in Table 1 along with the zero magnitude flux calibration. α Per and β Gem were the standard stars for these observations. The standard star calibrations derived from Johnson *et al.*⁹ are given in Table 2. The standard star observations were internally consistent to better than 0.05 mag and all satellite measurements were made within 0.05 air mass of a standard star measurement. 'Intermediate' air mass corrections were used⁸. Relative infrared reflectances were normalised to $R_\lambda(1.25\ \mu\text{m}) = 1.0$ after subtracting the appropriate solar colour. The solar magnitudes used are given in Table 2 and are derived from Low and Rieke⁸ and Labs and Neckle¹⁰. The log of satellite observations is listed in Table 3.

The results of the observations are given in Table 4 and illustrated in Fig. 1. The spectra clearly show an absorption feature centred between 3.1 and 3.4 μm . The band depth in the Ganymede spectrum is consistent with at least as much surficial coverage of water frost as estimated by Pilcher, *et al.*¹ and Kieffer and Smythe³. The absorption band in the spectrum of Callisto is shallower than that of Ganymede, but is of the same general shape. This indicates the presence of an optically significant amount of frost on Callisto and is consistent with a surface exposure of water frost which is somewhat less than on Ganymede.

Table 2 Stellar calibration

Wavelength (μm)	Sun*	Magnitude α Per†	β Gem†
1.25	-27.81	0.87	-0.49
2.22	-28.23	0.56	-1.09
3.03	-28.28	0.51	-1.14
3.12	-28.29	0.51	-1.14
3.43	-28.30	0.49	-1.16
3.45	-28.30	0.49	-1.16
3.73	-28.32	0.47	-1.18

*1.25- μm (J) magnitude calculated from the filter bandpass and the solar spectrum of Labs and Neckel¹⁰. 2.22 μm (K) from Low and Rieke⁸. All others are interpolated values from K and L magnitudes of Low and Rieke⁸.

†1.25- μm (J) and 2.2- μm (K) magnitudes from Johnson, *et al.*⁹. All others are interpolated values from K and L magnitudes of Johnson, *et al.*⁹.

This absorption feature is deeper than might be expected on both satellites. Assuming that the rest of the surface is covered by a 'grey' component (for example, silicate), Pilcher *et al.*¹ used the observed visual albedo and the observed depths of the 1.5- and 2.0- μm water frost features (upper limits for the case of Callisto) to determine the most probable surface coverage of water frost for Ganymede and Callisto. The depth of the 3.1- μm band can be predicted from the relative depths of the 1.5-, 2.0- and 3.1- μm water frost features, the estimated surface coverage of water frost, and albedo of the frost and 'grey' material. The predicted depth of the band is only about half as deep as observed on both Ganymede and Callisto (see Fig. 2a). One possibility is that the non-frost component has a flat relative reflectance from the visual through 2.5 μm and a lower reflectance in the 3-4- μm region. This may be indicative of the presence of dark hydrated minerals (for example, carbonaceous chondritic material) whose spectra show water absorption in the 3-4- μm region, but whose water absorption features shorter than 3 μm are suppressed by the presence of dark material (W. Salisbury, personal communication and ref. 11). Alternatively, the relative inconsistency between the band depths at 1.5 and 2.0 μm compared with 3.1 μm may be due to the intimate mixing of a significant amount of the 'observed' water frost with the 'grey' material, thus suppressing the short wavelength

Table 3 Log of Data

Date (UT)* (November 1976)	Wavelength (μm)	Magnitude	
		J III	J IV
7.34	1.25		4.36
7.35	3.0		5.17
7.35	3.1		5.53
7.36	3.4		5.49
7.36	3.45		5.04
7.36	3.7		4.96
7.41	3.7	4.80	
7.42	1.2	3.42	
7.43	3.0	5.41	
7.43	3.1	5.90	
7.43	3.4	6.34	
7.44	3.45	5.68	
29.28	3.45		5.05
29.28	3.7		4.99
29.29	2.2		4.16
29.30	1.2		4.43
29.32	3.4		5.44
29.32	3.0		5.04
29.33	3.1		5.64

*Each point consists of 2-3 separate measurements at the same wavelength taken less than 10 min apart.

bands relative to the much stronger 3.1- μm band. Either of these would be consistent with the models for the formation and evolution of the Galilean satellites that predict relatively unaltered surface crust of water frost and silicates for Callisto¹².

Recent airborne telescopic observations of the Galilean satellites and laboratory studies of hydrated minerals bear out both of these possibilities. The telescopic observations covering the spectral region 1-5 μm ($\Delta\lambda/\lambda \sim 0.03$) show a very strong absorption feature centred at about 2.8 to 3.0

Fig. 1 Relative reflectance of the satellites J III (Ganymede) (●) and J IV (Callisto) (○) normalised to $R_\lambda(1.25\ \mu\text{m}) = 1.0$. The broadband L point (3.45 μm) has been deleted for clarity.

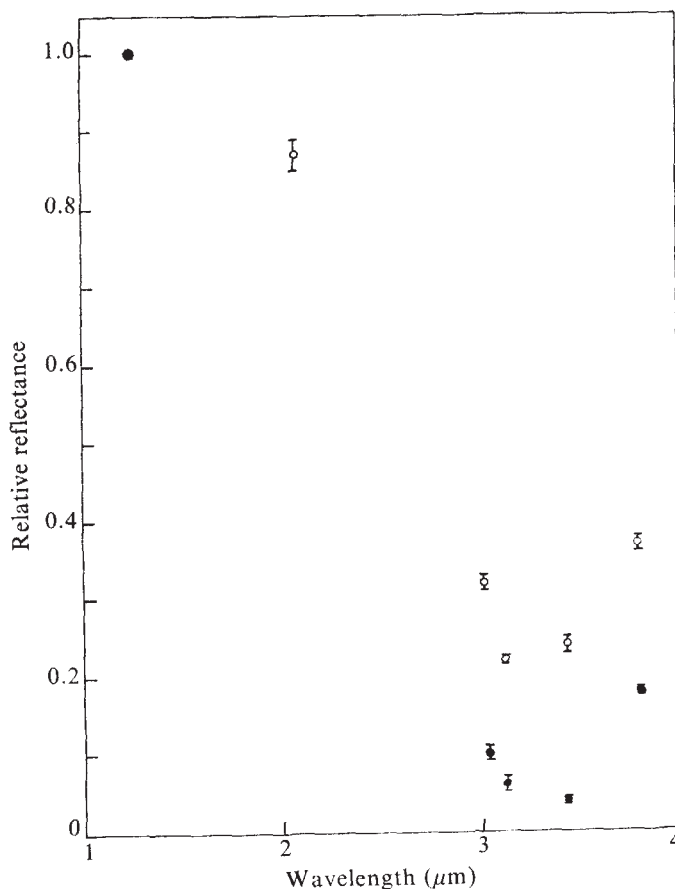


Table 4 Normalised reflectance

Wavelength (μm)	J III (6-7 November 1976)	J IV (6-7 November 1976)	J IV (28-29 November 1976)	J IV (Mean)
1.25	1.00	1.00	1.00	1.00
2.22			0.87 ± 0.02	0.87 ± 0.02
3.03	0.10 ± 0.01	0.31 ± 0.01	0.37 ± 0.02	0.32 ± 0.01
3.12	0.06 ± 0.01	0.22 ± 0.005	0.21 ± 0.01	0.22 ± 0.005
3.43	0.04 ± 0.005	0.22 ± 0.01	0.26 ± 0.01	0.24 ± 0.01
3.45	0.08 ± 0.005	0.34 ± 0.005	0.36 ± 0.01	0.34 ± 0.005
3.73	0.18 ± 0.005	0.37 ± 0.01	0.38 ± 0.01	0.37 ± 0.01

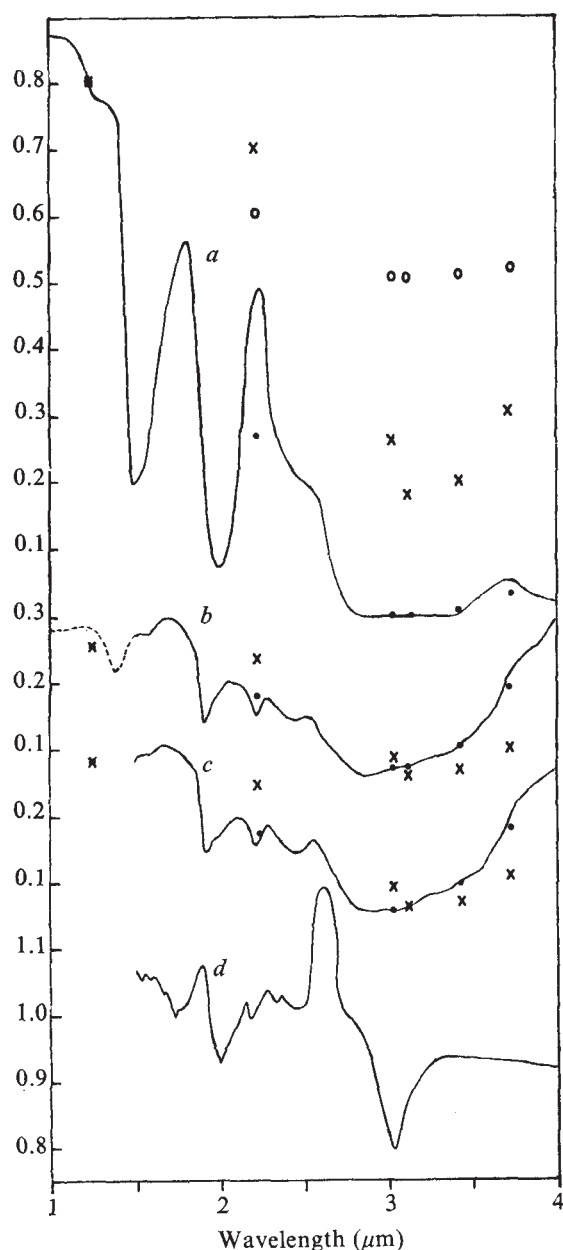


Fig. 2 Comparison of the relative spectral reflectance of Callisto to possible surface materials. *a*, Water frosts (absolute reflectance); *b*, montmorillonite, $(\text{Mg, Fe, Al})_x(\text{OH})_y$ ($\text{Si, Al, Fe})_x\text{O}_{20-n} \cdot n \cdot 6 \text{H}_2\text{O}$; *c*, Montmorillonite, cooled to about 150 K (reflectance, arbitrary units); *d*, Ratio *c/b*. For each curve, $\times$ and $\bullet$ represent the relative reflectances (for the filter band passes used) of Callisto and the materials in question, respectively. $\circ$ represent the relative reflectances in the filters of a mixture of water ice and a 'grey' material as estimated by Pilcher, *et al*¹ for the surface composition of Callisto (10% water frost and 90% 'grey' material). *a* is from Smythe⁴. All other curves are from Lebofsky¹⁵. The extension of curve *b* down to 1.0 μm was taken from the results of Nash and Fanale¹⁶.

μm which has been tentatively identified as being due to water of hydration¹³. These data are consistent with the results presented here, but are not of sufficient signal-to-noise to make an absolute match.

Using X-ray diffraction techniques¹⁴ it has been shown that when hydrated minerals (bentonite-water pastes in these studies) are frozen, all but two or three monomolecular layers of the interlamellar water migrates into the pore space to form ice. Reflectance studies (1.5–4.0 μm , $\Delta\lambda/\lambda \sim 0.002$) of room temperature and frozen ($\sim 150 \text{ K}$) montmorillonite and other minerals are being conducted¹⁵. Preliminary results are presented in Fig. 2, which shows that the water of hydration bands at 1.9 and 2.8 μm are decreased in strength and shifted to longer wavelengths due to the formation of water ice in the pore space. The absorption bands due to the water ice can be seen at 2.0 and 3.0–3.1 μm (Fig. 2*d*). It is clear from Fig. 2 that the hydrated minerals alone do not match the spectrum of Callisto because of the presence of bands at 1.4 and 1.9 μm and the differences in the 3–4- μm band shape, but it seems that this may be overcome by the addition of dark carbonaceous material (to suppress the 1.4- and 1.9- μm feature) and some water frost (to increase the band depth in the 3–4- μm region).

Based on the previous identification of water frost on the surface of Ganymede from shorter wavelength data, we infer the presence of water frost on the surface of Callisto from similar bands at 3–4 μm on both satellites. The depths of these bands may also indicate the presence of water of hydration, though how the frost and hydrated mineral may be mixed are as yet unknown.

I thank the University of Arizona, Lunar and Planetary Laboratory for telescope time and equipment, also Drs G. Rieke, F. Fanale, T. Johnson, D. Matson, and G. Veeder for helpful discussions. I was supported by a National Research Council Resident Research associateship during this research. This paper is JPL Planetary Publication 78-9 and was carried out under contract NAS 7-100, sponsored by the Planetary Geology Program Office, Office of Space Science, NASA.

LARRY A. LEBOFSKY*

Jet Propulsion Laboratory,
California Institute of Technology,
Pasadena, California 91103

Received 13 June; accepted 30 August 1977.

*Present address: Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721.

- Pilcher, C. B., Ridgeway, S. T. & McCord, T. B. *Science* **178**, 1087–1089 (1972).
- Fink, U., Dekkers, N. H. & Larson, H. P. *Astrophys. J.* **179**, L 155–159 (1973).
- Kieffer, H. H. & Smythe, W. D. *Icarus* **21**, 506–512 (1974).
- Smythe, W. D. *Icarus* **24**, 421–427 (1975).
- Irvine, W. M. & Pollack, J. B. *Icarus* **8**, 324–360 (1968).
- Gillett, F. C., Merrill, K. M. & Stein, W. A. *Astrophys. Lett.* **6**, 247–249 (1970).
- Hansen, O. L. *Icarus* **26**, 24–29 (1975).
- Low, F. J. & Rieke, G. H. *Meth. Exp. Phys.* **12A**, 415–462 (1974).
- Johnson, H. L., Mitchell, P. I., Iriarte, B. & Wisniewski, W. Z. *Commun. Lunar Planet. Lab.* **63**, 99–110 (1966).
- Labs, D. & Neckel, H. *Solar Phys.* **15**, 79–87 (1970).
- Johnson, T. V. & Fanale, F. P. *J. geophys. Res.* **78**, 8507–8518 (1973).
- Consolmagno, G. J. & Lewis, J. S. *Jupiter Studies of the Interior, Atmosphere, Magnetosphere and Satellites*, 1035–1051 (1976).
- Pollack, J. B., Witteborn, P., Erickson, E., Strecker, D., & Baldwin, H. presented at 8th Annual Planetary Geology Principal Investigator's Symposium, St. Louis, Missouri (1977).
- Anderson, D. M. & Hoekstra, P. *Am. Proc. Soil Soc.* **31**, 498–504 (1965).
- Lebofsky, L. A. presented at 8th A. Planet. Geol. Principal Investigator's Symp., St. Louis, Missouri (1977).
- Nash, D. B. & Fanale, F. P. *Icarus* **31**, 40–80 (1977).

Isotopic composition of uranium in chondritic meteorites

THE isotopic composition of uranium has been extensively measured for terrestrial U ores by precise gas mass spectrometry^{1,2}; measurements on other rock-types and minerals are limited³⁻⁵. Analyses of lunar material⁶⁻⁸ have shown the $^{238}\text{U}/^{235}\text{U}$ ratio to be the same, within experimental error, as the accepted value for terrestrial U (from pitchblende), namely 137.88 ± 0.14 (ref. 8). Measurements of the isotopic composition of U in five meteorites are reported here. The data show that variations in the U isotopic composition exist in these chondrites, and indicate the presence of components of low $^{238}\text{U}/^{235}\text{U}$ ratio. For meteorites there does not seem to have been any published attempt at a direct measurement of the $^{238}\text{U}/^{235}\text{U}$ ratio. In meteorite chronology it is generally assumed⁹⁻¹¹ that the U isotopic composition in meteorites is the same as that on Earth.

The samples used consisted of whole meteorite fragments free from fusion crust and broken from the interior of a given meteorite. A modified form of the method given by Arden and Gale¹² was used for the separation of U, which was carried out in clean-room conditions. The separated uranium was loaded, in water, on to an outgassed zone-refined V-shaped Re filament, previously completely covered on its top surface with a thin film of colloidal graphite and carefully evaporated

to dryness. Rigid control was maintained throughout on filament preparation, loading and running conditions. The mass spectrometer and ion-counting system used have been described¹².

The isotopic abundance of U, using NBS 950a standard, was measured in the range 0.2 to 20×10^{-9} g. The measured mean $^{238}\text{U}/^{235}\text{U}$ ratios ranged from 136.82 to 138.33 with a single-run internal precision of about $\pm 0.2\%$ ($1\sigma_m$). To check for any possible interference effects or suppression of emission, some runs were made after passing the U through the separation procedure and with an excess of tri-*n*-butyl phosphate. The variable mass fractionation between runs as measured by the standard deviation of replicate measurements was $\pm 0.37\%$.

The total range of U standard measurements about the mean value of 137.59 was $\pm 0.5\%$, which is similar to that obtained with the use of $^{233}\text{U}/^{238}\text{U}$ spike¹³ at the 0.1- μg U level. No measurable in-run fractionation, over the temperature range 1,600–1,700 °C, was observed. The U loaded on to the filament was in a highly purified state and no residue was visible. On the standard and sample runs a 'clean' spectrum was obtained. The metal-oxide ion ratio was of the order of 1,000:1 and the ^{238}U count rates were in the range 10^3 to 0.7×10^6 p.p.s. Sample data, measured U isotopic compositions, whole meteorite concentrations and computed overall whole meteorite isotopic compositions are given in Table 1.

The results show that U in the meteorites analysed is a mixture of components of widely different $^{238}\text{U}/^{235}\text{U}$ ratios. One

Table 1 Measured uranium isotopic compositions and concentrations

Meteorite	Mass (g)	Decomposition	Acid*	Measured isotopic composition† $^{238}\text{U}/^{235}\text{U}$	<i>m</i>	U‡ (10^{-9} g per g)	$^{238}\text{U}/^{235}\text{U}$ §
Parnallee LL3 (BM 34792)	1.62417	HF/HNO ₃	(a) HNO ₃	137.07 ± 0.26	46	10.60 ± 0.04	134.3
			(b) HCl	93.98 ± 0.34	27	0.468 ± 0.004	
			(c) HNO ₃	72.46 ± 0.85	17	0.020 ± 0.002	
Barwell L5 (BM 1966, 59)	1.58409	HF/HNO ₃	(a) HNO ₃	137.99 ± 0.36	27	12.38 ± 0.05	137.5
			(b) HCl	133.53 ± 1.04	17	0.702 ± 0.005	
			(c) HNO ₃	40.23 ± 0.73	6	0.010 ± 0.002	
Allende C3V (SI)	1	HF/HNO ₃	(a) HNO ₃	137.38 ± 0.30	32	15.01 ± 0.06	136.9
			(b) HCl	131.11 ± 0.37	20	0.410 ± 0.005	
			(c) HNO ₃	45.65 ± 1.13	22	0.020 ± 0.002	
	2	HF	(a) HCl	132.32 ± 0.34	32	5.86 ± 0.03	136.1
			(b) HNO ₃	139.11 ± 0.46	20	7.80 ± 0.03	
			(c) HCl	63.19 ± 0.19	20		
Bruderheim L6 (B.93 UA)	1.49691	HF/HNO ₃	(a) HNO ₃	136.12 ± 0.39	29	13.71 ± 0.07	136.6
			(b) HCl	126.21 ± 0.89	20		
Richardton H5 (100 h ASU)	1	HF/HNO ₃	(a) HNO ₃	137.14 ± 0.18	30	7.47 ± 0.03	136.6
			(b) HCl	122.26 ± 0.79	23	0.249 ± 0.003	
	2	HF	(a) HNO ₃	123.51 ± 0.21	31	9.37 ± 0.04	124.4
			(b) HCl	135.77 ± 0.36	22	0.779 ± 0.005	
	3	HF	(a) HCl	134.66 ± 0.38	16	4.27 ± 0.02	106.8
			(b) HNO ₃	92.07 ± 0.12	40	5.58 ± 0.02	
	4	HF/HNO ₃	(a) HNO ₃	136.71 ± 0.14	15	8.76 ± 0.04	133.1
			(b) HCl	103.72 ± 0.14	39	0.808 ± 0.006	
	5	HF/HNO ₃	(a) HNO ₃	136.50 ± 0.28	21	9.72 ± 0.05	136.5
			(b) HCl	136.58 ± 0.21	38	0.96 ± 0.01	
(BM 1920, 511)	6	HF/HNO ₃	(a) HNO ₃	134.98 ± 0.36	19	10.22 ± 0.06	134.6
			(b) HCl	132.51 ± 0.43	29	1.54 ± 0.02	
		HF/HNO ₃	(a) HCl	135.68 ± 0.22	34	10.08 ± 0.04	135.6
			(b) HNO ₃	132.10 ± 0.55	23	0.228 ± 0.004	

Sources of Meteorites: ASU, Arizona State University (Centre for Meteorite Studies); UA, University of Alberta; SI, Smithsonian Institution; BM, British Museum (Natural History).

* (a), Dried residue heated, in autoclave for 1–2 d at $\sim 180^\circ\text{C}$ with 7 ml of either 7M HNO₃ or 6M HCl. The solution was removed, a further similar volume of the acid added and the heating cycle repeated. (b), As for (a), with next acid. (c), A very small residue often remained after (a) and (b). For some of the samples this residue was heated with 16M HNO₃ or 6M HCl at $\sim 180^\circ\text{C}$ for 4 d.

† Measured ratios were corrected for dead-time but not for blank. (Complete blank ~ 0.01 ng, separation blank ~ 0.004 ng.). *m*, No. of data cycles (each cycle consists of the mean of 5 measured ratios). σ_m , Standard deviation of means.

‡ Calculated using measured $^{238}\text{U}/^{235}\text{U}$ ratio, but not corrected for blank.

§ Overall $^{238}\text{U}/^{235}\text{U}$ ratio for the whole meteorite sample—computed from the total no. of atoms of the respective isotopes in each extraction.

component of terrestrial, or near terrestrial, isotopic composition is usually present at the normal abundance level (~ 10 parts in 10^9) and is readily dissolved (after HF attack) by HCl or HNO_3 . The others with lower measured $^{238}\text{U}/^{235}\text{U}$ ratios are present in low abundance in the whole meteorite, and, as indicated by the prolonged acid attack at elevated temperature and pressure required to release the components (see Table 1), seem to be contained in very resistant phases. The distribution of the latter components within the whole meteorite samples is probably very heterogeneous. Fission track studies in chondrites show that changes in U concentration by factors of 10^5 may be observed over distances of $\sim 10^{-2}$ cm (ref. 14).

The dissolution process, however, can influence the contribution made by each component to the U isotopic composition in each fraction. After decomposition, the dried complex fluoride residue is leached with a high concentration of acid at elevated temperature and pressure. This dynamic equilibrium process releases species occluded in the precipitate so that they remain in the solution phase, and has been shown to give quantitative solution of U ('major component') in HCl (ref. 12) and HNO_3 . If, however, one or more component phases are present which are attacked by HNO_3 or HCl but possibly only to a very limited extent (in these experimental conditions) by HF or HF/ HNO_3 during sample decomposition, then the rate and extent of dissolution of this phase depends on several factors. Among these are its physical state (dispersed surface coating or discrete mineral grains) actual time and length of exposure to the acid during the leaching process, as well as the activity of the acid in solutions of varying ionic strength. (The data obtained for the U standards, as well as the detailed analytical procedure used may be obtained from the author.)

Some possible mechanisms by which heterogeneities could arise are: (1) In the two-spike model, Wasserburg *et al.*¹⁵ predicted that the *r*-process nuclei occurred in three stages; a large amount of initial production early in the history of the galaxy, followed by a relatively quiescent period ($\sim 3 \times 10^9$ yr) terminated by a nucleosynthetic event that possibly initiated the separation of the solar system. Incomplete mixing of matter from the two events in the relatively short time interval between the predicted terminal spike at 4.8 to 4.6 Gyr and the formation of the solar system could occur: matter from the last nucleosynthetic event would have a low $^{238}\text{U}/^{235}\text{U}$ ratio. (2) The presence of grains in the solar nebula that had previous existence in the interstellar medium could, if incorporated into the meteorite parent body, result in a variable but low $^{238}\text{U}/^{235}\text{U}$ ratio¹⁶. (3) Grains with a low $^{238}\text{U}/^{235}\text{U}$ ratio could result from the explosion of a Type II supernova which could have triggered the collapse of a nearby interstellar cloud and led to the formation of the solar system¹⁷. (4) Chemical fractionation of Cm and U in some region of the early solar system. The nuclide ^{247}Cm decays to ^{235}U ; thus if a large $^{247}\text{Cm}/\text{U}$ fractionation occurred it would have produced enhancement in the abundance of ^{235}U , resulting in a low $^{238}\text{U}/^{235}\text{U}$ ratio¹⁸.

I thank M. C. B. Abranches for evaluation of the mass spectrometry of U. This study was supported by a grant from the NERC.

JOHN W. ARDEN

Department of Geology and Mineralogy,
University of Oxford,
Parks Road, Oxford, UK

Received 4 August; accepted 1 September 1977.

1. Cowan, G. A. & Adler, H. H. *Geochim. cosmochim. Acta* **40**, 1487 (1976).
2. Hamer, A. N. & Robbins, E. J. *Geochim. cosmochim. Acta* **19**, 143 (1960).
3. Doe, B. R. & Newell, M. F. *Am. Mineral* **50**, 613 (1965).
4. Rosholt, J. N., Harshman, E. N., Shields, W. R. & Garner, E. L. *Econ. Geol.* **59**, 570 (1964).
5. Rosholt, J. N., Garner, E. L. & Shields, W. R. *U.S. Geol. Survey Prof. Paper* 501-B, B84-B87 (1964).
6. Rosholt, J. N. & Tatsumoto, M. *Geochim. cosmochim. Acta Suppl.* **1**, 1499 (1970).
7. Fields, P. R., Diamond, H., Metta, D. N., Stevens, C. M. & Rokop, D. J. *Geochim. cosmochim. Acta Suppl.* **2**, 1571 (1971).
8. Barnes, I. L., *et al.* *Geochim. cosmochim. Acta Suppl.* **3**, 1465 (1972).
9. Tatsumoto, M., Unruh, D. M. & Desborough, G. A. *Geochim. cosmochim. Acta* **40**, 617 (1976).
10. Chen, J. H. & Tilton, G. R. *Geochim. cosmochim. Acta* **40**, 635 (1976).

11. Hutchison, R., Gale, N. H. & Arden, J. W. *Nature* **254**, 678 (1975).
12. Arden, J. W. & Gale, N. H. *Anal. Chem.* **46**, 687 (1974).
13. Ridley, R. G., Daly, N. R. & Dean, M. H. *Nucl. Instr. Meth.* **34**, 163 (1965).
14. Fleischer, R. L. *Geochim. cosmochim. Acta* **32**, 989 (1968).
15. Wasserburg, G. J., Schramm, D. N. & Hüneke, J. C. *Astrophys. J.* **157**, L91 (1969).
16. Clayton, D. D. *Earth planet. Sci. Lett.* **35**, 398 (1977).
17. Cameron, A. G. W. & Truran, J. W. *Icarus* **30**, 447 (1977).
18. Blake, J. B. & Schramm, D. N. *Nature phys., Sci.* **243**, 138 (1973).

Stacking faults in dolomite

DOLOMITE, an important rock-forming mineral, is a mixed carbonate possessing trigonal symmetry ($R\bar{3}$), with alternate layers of Mg and Ca atoms separated by planes containing the carbonate groups. Its ordered structure offers the possibility of extended planar faults of low energy and Barber¹ has shown evidence for stacking disorder in dolomite of high metamorphic grade. There is no previous evidence, however, for the creation of extended faults as a direct result of deformation processes. We have prepared single crystals, approximately 25 mm² in cross section and 10 mm long, and tested them in compression at temperatures up to 800 °C while subjected to a confining pressure of either 3 or 7 kbar. Specimens were mostly oriented to give a large resolved shear stress on the basal planes while minimising the stress for mechanical twinning on the $f = \{01\bar{1}2\}$ planes. (We use four-digit indices for planes and directions referred to the structural hexagonal unit cell, with $c = 16.01$ Å and $a = 4.81$ Å.) It has been shown (optically)² that slip on (0001) and twinning on $\{01\bar{1}2\}$ are common modes of deformation on dolomite. Using TEM methods, Barber¹ has shown that slip on $\{01\bar{1}2\}$ can also occur in dolomite. The current observations were also made by high voltage transmission electron microscopy, using sections which were thinned first mechanically and then by ion bombardment.

The extended planar faults are generated by slip on the basal planes and when viewed with the electron beam aligned parallel to a $\langle 2\bar{1}\bar{1}0 \rangle$ direction, they are visible edge-on as very fine lines of low contrast. They are thus easily distinguishable from long screw-type basal dislocations which are profuse in most samples. By tilting a foil of known orientation, ($2\bar{1}\bar{1}0$), through a measured angle about the $[01\bar{1}0]$ zone axis (the line of intersection of a fault with the foil plane) and knowing the foil thickness, the fault plane has been shown to be unequivocally (0001). Local specimen thickness was determined easily because of the presence of small cleavage cracks on $\{10\bar{1}4\}$.

Figure 1a shows the faults in minimum contrast, using a 0330 reflection; the dislocations terminating the faults are clearly visible. The faults also have minimum contrast for reciprocal lattice vector $g = 1\bar{1}20$. In Fig. 1b, the beam is still nearly parallel to $[2\bar{1}\bar{1}0]$, but both $01\bar{1}5$ and $02\bar{2}10$ reflections are operating to give strong fault contrast. In Fig. 1c a tilt of approximately 20° is applied to show fringed fault contrast, using $+g = 1\bar{1}08$ and $-g$, while in Fig. 1d the fault is invisible ($g = 2\bar{1}1\bar{1}8$) and the partial dislocations are in contrast. Although these observations were taken at 1,000 kV and dolomite is elastically anisotropic, the clarity of these contrast changes strongly suggests that we can calculate the displacement vector R in the usual manner, taking $g \cdot R = n$, where n must be zero or an integer for zero contrast with a given reciprocal lattice vector g . Using the above data, this gives as a likely value for R the vector $\frac{1}{3}[1\bar{1}00]$, which is not a proper translation vector of the dolomite structure. As might be anticipated, the vector lies in the basal plane and, for example, when applied to a magnesium atom, translates it at 30° to the normal $[1\bar{2}\bar{1}0]$ slip vector to a position directly above a calcium atom. This creates a stacking fault in the structure (not an antiphase boundary); the stacking fault is removed with a further displacement of $\frac{1}{3}[01\bar{1}0]$, giving

$$\frac{1}{3}[1\bar{1}00] + \frac{1}{3}[01\bar{1}0] = \frac{1}{3}[1\bar{2}\bar{1}0]$$

which is the normal slip vector of length 4.81 Å. This reaction, of course, identifies the observed partial dislocations. We have

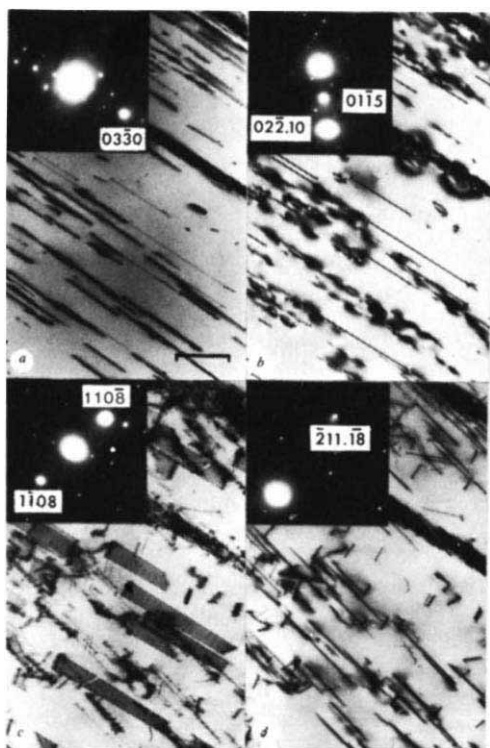


Fig. 1 Basal stacking faults in dolomite single crystal 4071 XD, compressed at 500 °C under a confining pressure of 3 kbar to a strain of 2.5% at a strain rate of 10^{-5} s^{-1} . The foil plane is (2110) and the faults are imaged at 1 MeV in various diffraction conditions: a, 0330 reflection; scale bar, 1 μm ; b, 0115 and 022.10 reflections; c, 1108 and 1108 reflections; d, 211.18 reflection.

found that such stacking faults occur extensively in dolomite crystals compressed at temperatures of 300–500 °C.

We have also observed that faint fringe contrast is sometimes created by the passage of dislocations on the {0112} slip planes. We have failed, however, to find a diffraction vector which brings the fringes into strong contrast and the fringes do not always terminate at dislocations but sometimes extend throughout the active part of the slip plane. It thus seems that the faint fringes, which indicate residual elastic strain, may be associated with a small fractional displacement vector, as occurs in orthopyroxene³.

Evidence for stacking faults generated by slip in carbonates is new and it further emphasises the differences in mechanical behaviour between dolomite and calcite⁴. Until the work of Turner and Orozco⁵ on metamorphic calcite there was no firm evidence for basal slip in calcite, although it is the favoured low-temperature mechanism in dolomite. Brailon and Serughetti⁶ have reported small (0001) planar faults in calcite but have attributed them to growth twinning. Our results, which demonstrate the existence of extensive stacking faults in dolomite, together with partial dislocations, are therefore novel and important.

The work was supported by SRC grant GR/A 14599. We thank E. Amundarein, Magnesitas-Navarra, for dolomite crystals.

D. J. BARBER

Physics Department,
University of Essex,
Wivenhoe Park,
Colchester, Essex, UK

H. C. HEARD

Lawrence Livermore Laboratory,
University of California,
P.O. Box 808, Livermore,
California 94550

M. S. PATERSON

Research School of Earth Sciences,
Australian National University,
Canberra, A.C.T.,
Australia

H. R. WENK

Department of Geology and Geophysics,
University of California,
Berkeley,
California 94720

Received 8 August; accepted 2 September 1977.

1. Barber, D. J. *Tectonophysics* **39**, 193 (1977).
2. Higgs, D. V. & Handin, J. *Bull. Geol. Soc. Am.* **70**, 245 (1959).
3. McLaren, A. C. & Etheridge, M. A. *Contrib. Mineral. Petrol.* **57**, 163 (1976).
4. Barber, D. J. & Wenk, H. R. *Tectonophysics* (in the press).
5. Turner, F. J. & Orozco, M. *Contrib. Mineral. Petrol.* **57**, 83 (1976).
6. Brailon, P. & Serughetti, J. *Phys. Stat. Sol.* **33**, 405 (1976).

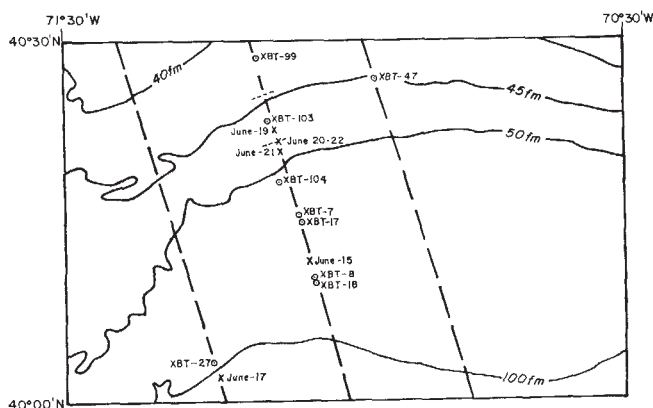
Acoustic imaging of the New England shelf-slope water mass interfaces

PULSED high frequency acoustics has been shown to have unique applications in the detection and measurement of oceanic phenomena such as internal waves^{1–3} and dispersion of ocean-dumped sewage sludge^{4,5} or dredge material⁶. With moderate to high pulse repetition rates, a virtually continuous record of water column scattering strength is available to depths of the order of 500 m from a ship travelling at speeds of 4 or 5 knots. We used this feature to detect acoustically the interfaces associated with an intrusion of cold New England shelf water into the warmer slope water during a 9-d period in June 1976.

From 15 to 23 June 1976 measurements of effects of internal waves were made from the National Oceanic and Atmospheric Administration (NOAA) ship George B. Kelez in an area on the New England continental shelf partly shown in Fig. 1. Although measurements were made on three tracklines shown, most of the work concentrated on the centre trackline. A 1-kW 20-kHz pulsed acoustic system, specially modified for detection of weak signals scattered from the water column, was operated continuously throughout the cruise. Field operation of this system has been described previously^{4,7}. Profiles using expendable bathythermographs (XBT) and a conductivity–temperature–depth probe (CTD) were obtained at regular intervals.

Throughout the cruise, the acoustic record revealed a reflecting layer (Fig. 2) of limited horizontal extent and

Fig. 1 Region of New England continental shelf where water mass interfaces were observed. Contours are in fathoms, dashed lines are tracklines along which most work was performed. X with a date denotes an observation of the intersection of the scattering layer offshore boundary and the ship's trackline. The light dashed lines represent the width of the intrusion observed on 22 June.



rather variable location. Some observations of the offshore boundary of this layer are indicated in Fig. 1. The layer was observed in all cases (18 observations) to have a slope opposite to that of the bottom. As shown along the ship's trackline, the layer displayed apparent widths of 2.8–7.4 km. The last few records (for example, Fig. 2) of the scattering layer also show a second layer higher in the water column and parallel to the surface, but apparently descending to join the ascending lower layer at their offshore boundary. This made a sufficiently dramatic feature to finally get our attention. We then obtained XBT and CTD profiles through the feature when the cruise plan permitted.

A series of three XBT profiles (Fig. 3) across the layers revealed that they occur in a region of transition between two distinct water masses. XBT 99 shows the bottom shelf water to have a temperature of about 7 °C while XBT 104 shows the minimum temperature of the deeper slope water to be about 10 °C. XBT 103, taken in the region where the layers were present shows interleaving of 7 °C water with 10 °C water. The depths of the upper and lower steep

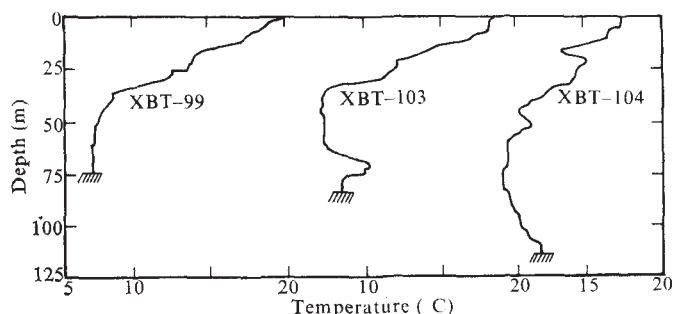


Fig. 3 Three XBT profiles (locations shown in Fig. 1) showing the changing make-up of the water column across the region in which acoustic scattering layers were identified as shelf-slope water mass interfaces. The temperature scale is actually three scales joined at the 20 °C points.

water interface to be further south. We observed no reflecting layer associated with the water-mass interface on the eastern trackline but found the transition took place between XBT 46 42° 32'N and XBT 47 on 18 June. It would seem, then, that the shelf-slope water transition zone was in a north east-south west direction in the vicinity of this experiment. Curiously, from 19 through 22 June the deep water boundary of the transition zone was roughly stationary as shown in Fig. 1. Thus the boundary moved 18 km along the centre trackline from 16 to 19 June and moved very little in the next four days.

The shelf-slope water transition zone apparently fluctuates significantly in location on a time scale of a few days. Also, Cresswell¹⁰ has observed distinct parcels of shelf water in the deeper slope water indicative of a mixing process. If the conditions for scattering at these water mass interfaces are often suitable for acoustic scattering, it may be possible to observe acoustically the shelf-slope water exchange process occurring in the summer on the New England continental shelf. XBT and CTD profiles would be needed only occasionally to confirm that the scattering layers are the interfaces of interest. Most of the ship time would then be spent covering the area of interest in a regular fashion with sufficient frequency to resolve water mass movements as they occur.

FRED C. NEWMAN
JOHN R. PRONI
DONALD J. WALTER

Sea-Air Interaction Laboratory,
Atlantic Oceanographic and Meteorological Laboratories,
Environmental Research Laboratories,
National Oceanic and Atmospheric Administration,
Miami, Florida, 33149

H. MICHAEL BYRNE

Pacific Marine Environmental Laboratory,
Environmental Research Laboratories,
National Oceanic and Atmospheric Administration,
Seattle, Washington, 98105

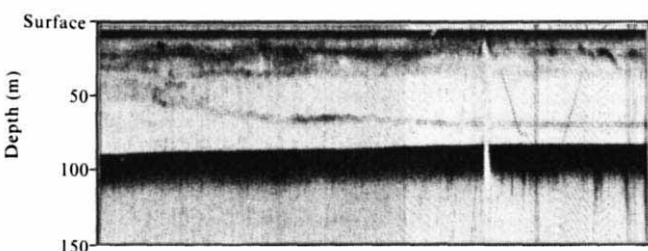


Fig. 2 A portion of the acoustic image showing the scattering layers associated with the intrusion interfaces (frequency 20 kHz). The left portion of the image, with darker background, represents a distance of 5.6 km. The portion to the right of the interrupted bottom signal was made with the ship on station for a CTD profile. The CTD echo signal is visible on the image.

temperature gradients in XBT 103 coincide with the depths of the acoustically reflecting layers measured at the moment the XBT 103 profile was taken. The gradient depths of XBT 103 (taken at 0729 GMT, 22 June) may be compared with the depths of the reflecting layers in Fig. 2, observed at 1325 GMT, 22 June. The slightly later acoustic image was chosen for its higher quality and clearer manifestation of the layers joining at the offshore boundary. In the right hand portion of Fig. 3 (the region of lighter background), the Kelez took station and a CTD profile was made and can be seen falling and then rising through the layers. The CTD temperature profile was similar to XBT 103 but showed very much steeper temperature gradients at the top and bottom of the 7 °C intrusion.

The horizontal transition from 7 to 10 °C water is a familiar feature. It is well known⁹ that by June the surface shelf water has warmed due to increased net solar radiation while the bottom shelf water remains cold, typically about 7 °C. The slope water is not cooled as much during winter mixing and has a minimum temperature of about 10 °C. During winter the shelf-slope water front is a well formed feature⁹, but spring surface warming over the shelf and slope removes the horizontal temperature contrast except in the lower water column. The reflecting layers manifested in the acoustic image are associated with nearly horizontal interfaces formed by interleaving of the shelf and slope water masses during spring and summer.

The location of the shelf-slope water interface on the centre trackline migrated about 18 km northwards during the nine-day span of observations. Reflecting layers observed on 15 and 16 June were confirmed to be shelf-slope water interfaces by XBT pairs (7, 8, and 17, 18 respectively) taken north and south of the observed layers. On 17 June a layer observation on the western trackline, confirmed by XBT 26 (not shown on map) and XBT 27, showed the shelf-slope

Received 26 July; accepted 14 September 1977.

1. Weston, D. E., Review, J., Jones, F. R. & Ranster, J. W. *Rep. ARL/L/R88* (Admiralty Res. Lab., Teddington, 1970).
2. Proni, J. R. & Apel, J. R. *J. geophys. Res.* **80**, 1147–1151 (1975).
3. Fischer, F. H. & Squier, E. D. *J. acoust. Soc. Am.* **58**, 1315–1317 (1975).
4. Proni, J. R., Newman, F. C., Sellers, R. L. & Parker, C. *Science* **193**, 1005–1007 (1976).
5. Proni, J. R. *et al. J. geophys. Res.* (submitted).
6. Proni, J. R., Rona, D. C., Lauter, C. A. Jr & Sellers, R. L. *Nature* **254**, 413–415 (1975).
7. Proni, J. R. *et al. Deep Sea Res.* **23**, 831–837 (1976).
8. Beardsley, R. C. & Flagg, C. N. *Mem. Soc. R. Sci. Liège* **6**, 209–225 (1976).
9. Wright, W. R. *J. mar. Res.* **34**, 1–14 (1976).
10. Cresswell, G. M. *Rep.* 67–35 (Woods Hole Oceanogr. Inst., Massachusetts, 1967).

Effect of climate on chemical composition of fossil bones

THE suggestion that microanalysis of nitrogen, fluorine and uranium in fossil bones should be used for age determination¹⁻⁴ assumes that the concentration of these elements varies uniformly in time. Vonach⁵ has plotted calibration curves for the N and F content of fossil bones from the past 10⁵ yr. The nitrogen content, N , can be approximated by

$$\log N = -0.135 \log t + 0.681$$

where N is weight per cent and t is age in years. The chemical changes that occur in buried bones depend on the composition of the surrounding minerals and the environmental conditions⁶. The concentration of a given element in bone is also influenced by biochemical differences between samples. Consequently, large fluctuations in N can be expected around the gross trend described by Vonach. We report here a fluctuation of the N content of bones that seems to correlate with climatic change.

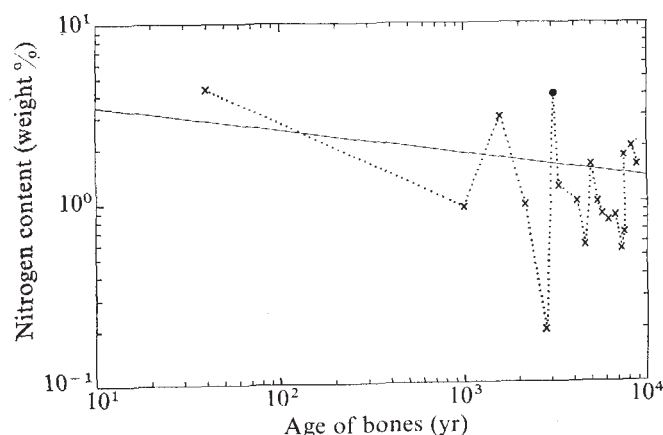


Fig. 1 N content of bones as a function of absolute age. . . . x . . . , Our results; ●, at 3,100 yr and —, from ref. 5.

We investigated the concentration of N, Fe, Al and Ca in bones from the past 9,000 yr. We chose this time interval because of the relatively large number of dated samples found and the significant climatic change during the interglacial phase. We used human vertebrae of known ages buried in the Great Hungarian Plain.

The N, Fe and Al content was determined by fast and thermal neutron activation analyses using the $^{14}\text{N}(n,2n)^{13}\text{N}$, $^{56}\text{Fe}(n,p)^{56}\text{Mn}$ and $^{27}\text{Al}(n,\gamma)^{28}\text{Al}$ reactions, respectively. The

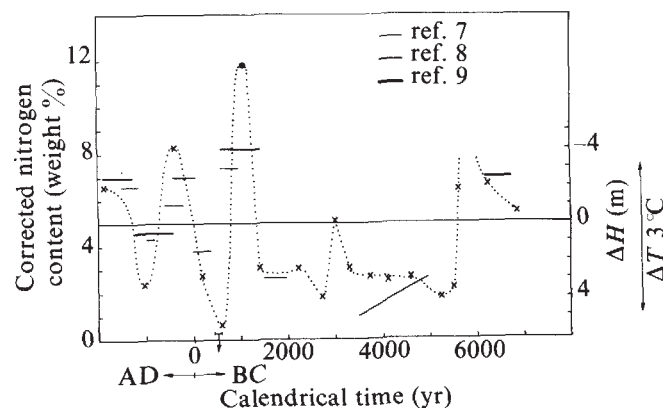


Fig. 2 Corrected N content of bones (. . . x . . .) and the sea level difference (straight lines) as a function of calendrical time.

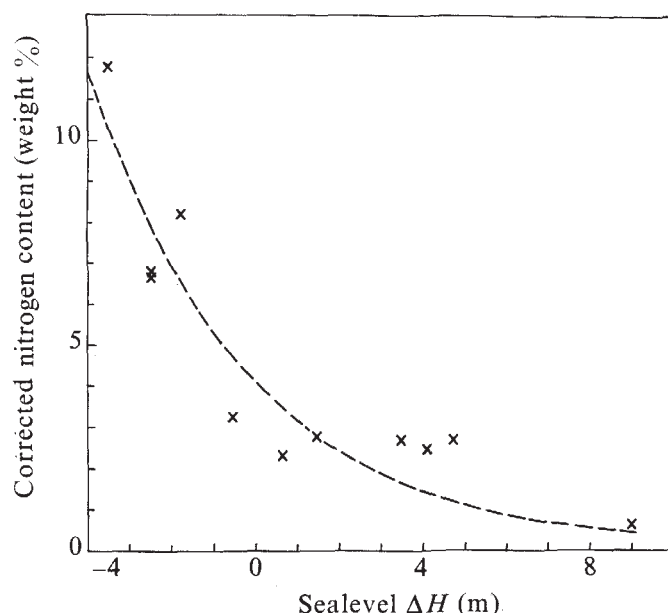


Fig. 3 Corrected N content of bones as a function of sea-level difference. $N = 4.1405 \exp(-0.26\Delta H)$.

concentration of Ca and Fe was measured by X-ray fluorescence using an Si(Li) detector and ^{55}Fe and ^{125}I exciting sources. Figure 1 shows N content as a function of the age of the bone. There is a fine structure (dotted curve) superimposed on the Vonach's gross trend (continuous curve).

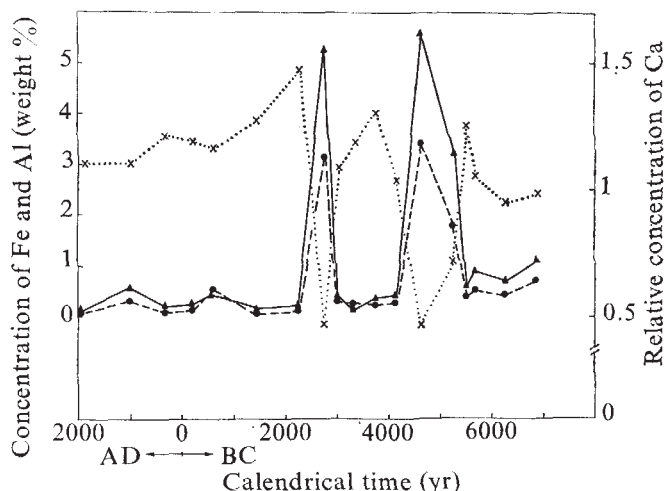


Fig. 4 Concentration of Al (▲), Fe (●) and Ca (x) in bones as a function of calendrical time.

The fluctuations in the data cannot be explained by biochemical differences between individuals and the uncertainty in the determination of N, for which the error does not exceed $\pm 10\%$. Results of *in vivo* neutron activation analysis of patients have shown that deviations in N content are less than $\pm 30\%$ even for different diseases^{12,13}.

There seems to be a periodicity of about 2,000 yr between the minima of the fluctuations. Accepting the gross trend for N content, we normalised the measured data to the date of burial. It should be noted, however, that the corrected N contents in some cases are abnormally larger than those obtained for modern bones. These normalised values and the sea-level differences, ΔH , from the present-day value⁷⁻⁹ are given in Fig. 2 as a function of calendrical time. An extremely high value of sea level (9 m) was given by Brooks¹¹ for 500 BC. Figure 2 shows a correlation between N content and sea level. The correlation can be approximated

by the following relationship: $N = 4.1405 \exp(-0.26\Delta H)$ (Fig. 3). Brooks has pointed out a linear correlation between climate and sea level. On the basis of the data for sea level⁷ and temperature⁹ for the interval from 1,000 BC to 5,000 BC, a value of $\Delta H/\Delta T \sim 2.6$ m per °C can be estimated. It can be concluded that the minima in N content indicated in Fig. 1 coincide fairly well with the climatic maxima. It seems to be a correlation between the 'neo-glacial cycle'¹⁰ and the N content of bones. The deviations from the predicted solid line in Fig. 1 can be attributed to the change in the climate of the age from which the bones originate.

The measured values for Al, Ca and Fe content of human bones as a function of calendrical time are given in Fig. 4. Both the trends and the absolute values for iron measured by X-ray fluorescence and neutron activation are in good agreement within $\pm 5\%$, proving the reliability of the methods applied. Figure 4 shows changes in the concentrations of all the elements investigated.

An opposite trend is shown for Fe and Al on the one hand, and Ca and N on the other hand. Chemical separation showed that the Fe and Al contents belong to the organic part of the bone.

We thank Professor J. Nemeskéry for discussions and for bone samples.

Cs. M. BUCZKO

Isotope Laboratory

L. VAS

Institute of Experimental Physics,
Kossuth University,
Debrecen, Hungary

Received 28 March; accepted 11 August 1977.

1. Middleton, J. *Proc. Geol. Soc. Lond.* **4**, 431–433 (1844).
2. Oakley, K. P. in *Anthropology Today* (Kroeber, Chicago, 1953).
3. Oakley, K. P. in *The Scientist and Archaeology*, (ed. Pyddoke, E.) (Oxford University Press, London, 1963).
4. Oakley, K. P. in *Science in Archaeology* (eds Brothwell, D. R. & Higgs, E. S.) (Praeger, London, 1963).
5. Vonach, V. *Atomki Közlemények* **18**, 247–258 (1976).
6. Protisch, R., *Ninth Congress Union Internationale des Sciences Préhistoriques et Protohistoriques*, Nice, 180–211 (1976).
7. Fairbridge, R. W. *Physics and Chemistry of the Earth* **4**, 99–185 (1961).
8. Komlódi, J. M. *Természet Világa* **12**, 535–540 (1971).
9. Ambrózy, P., Czelnai, R. & Götz, G. *Fizikai Szemle* **27**, 54–62 (1977).
10. *IAEA Bulletin* **19**, 32–35 (1977).
11. Brooks, C. E. P. *Climate Through the Ages* (London, 1950).
12. Cohn, S. H., Dombrowski, C. S., Hauser, W., Klopfer, I. & Atkins, H. L. J. *clin. Endocrinology Metabolism* **33**, 719–728 (1971).
13. Cohn, S. H., Cinque, T. J., Dombrowski, C. S. & Letteri, J. M. J. *Lab. clin. Med.* **79**, 978–994 (1972).

Physiological exclusion of magnesium from *Mytilus edulis* calcite

CALCITE produced by many marine organisms (including molluscs¹, brachiopods², coccolithophorids³, and planktonic foraminifera^{4,5}) contains only about 5% as much magnesium as is predicted by laboratory studies of distribution coefficients^{6,7} (ratio of the Mg:Ca ratio in the solid against that in the solution). Understanding the cause of this low magnesium content is important to our comprehension of the calcification process. It is also of considerable interest to the field of low temperature geochemistry for several reasons. First, calcite is the major authigenic component of marine sediments, and its magnesium content is an important feature of its bulk composition. Second, the controls on the magnesium content of invertebrate skeletons must be understood if Mg:Ca ratios of calcitic fossils are to be used as indicators of sea salt palaeochemistry⁷ or sea surface palaeotemperature⁸. Third, the mechanism of magnesium inclusion may determine the role of magnesium in regulating selective dissolution of the tests from planktonic foraminifera⁴. Fourth, since calcite solubility depends on the magnesium content of the calcite, the mode of magnesium inclusion is a factor to be considered in calculating the degree of carbonate saturation in the

oceans. Finally, factors governing the mode of magnesium incorporation into skeletal calcite are important to sedimentary petrologists using magnesium as a diagenetic indicator. We report here that the mussel *Mytilus edulis*, which normally secretes low magnesium calcite, incorporates anomalously high amounts of magnesium into its calcitic shell layer when it is grown in solutions of higher than normal magnesium content. Thus in normal conditions *Mytilus* physiologically excludes magnesium from its shell-forming fluid, and at higher magnesium concentrations ion regulatory systems breakdown, causing a substantial increase in the amount of magnesium coprecipitated into *Mytilus* shell calcite.

Young *M. edulis* (7–15 mm long) were collected from Narragansett Bay, Rhode Island and cultured for 8 weeks in aerated 4.5-l glass aquaria. The water was changed at weekly intervals and the mussels were fed a suspension of *Isochrysis*. Twenty mussels were placed in each tank. The culture solutions were prepared by mixing one part natural seawater with one part synthetic 'seawater'. A range of Mg:Ca ratios in the culture solutions was obtained by changing the proportions of sodium and magnesium in the synthetic 'seawater' solution. A constant total ionic strength (0.64 ± 0.04 M) could then be maintained between different composition solutions while also maintaining a constant calcium concentration similar to that of seawater (9.3 mM) in all the solutions except the lowest Mg:Ca ratio solution. To make the lowest Mg:Ca ratio solution (Mg:Ca = 0.75) a calcium concentration of 33 mM was used in the culture solution.

M. edulis forms both calcitic and aragonitic shell layers. Calcite samples of shell grown during this experiment were collected from the newly grown calcitic-shell edges outside the pallial line. These samples were analysed by atomic absorption spectrophotometry for magnesium and calcium^{5,9}. Ratios of the molar concentrations in both the solution and shell are reported since the ratios more clearly show the relative fractionation effects between the two ions. The growth of each mussel was recorded by measuring its length at the beginning and end of the experiment.

The Mg:Ca atom ratios of the newly grown shell are compared with those of the culture solutions in Fig. 1. Normal seawater has a Mg:Ca ratio of 5.2. Mussels grown in artificial media with the same Mg:Ca ratio as natural seawater produce the same shell Mg:Ca ratio as a control set of mussels grown in 100% natural seawater. Mussels grown in solutions with a Mg:Ca ratio less than 7 show a gradual increase in the shell Mg:Ca ratio with increases in the solution Mg:Ca ratio, except for the lowest solution ratio sample which has a slightly elevated shell ratio. As the Mg:Ca ratio of the solution increases above 7 there is a dramatic increase in the amount of magnesium incorporated into the shell.

Inorganic distribution coefficient data cannot account for the Mg:Ca ratios of the shells. Distribution coefficient data determined by Katz⁷ for low Mg:Ca ratio solutions predict that calcite precipitated from a solution with a Mg:Ca = 0.75 (the lowest ratio tested) will have a calcite Mg:Ca ratio of 40×10^{-3} ; this is 20 times more magnesium than we find in *Mytilus* shells grown in these conditions. Studies of the inorganic precipitation of calcite from seawater reveal Mg calcites containing >10 mol per cent MgCO₃ (ref. 6), although Berner⁶ suggests that 2–7 mol per cent MgCO₃ is the most stable composition in the presence of seawater. Distribution coefficient data are not available for solution Mg:Ca ratios of >5. But, surface-solution equilibrium exchange data show that near a Mg:Ca solution ratio of five surface distribution coefficients decrease¹⁰; if this is any guide, we would also expect bulk distribution coefficients to decrease. The increase in mag-

nesium observed in our experiments is directly opposite to this expectation and suggests that inorganic processes are not responsible for regulating the magnesium content of *Mytilus* calcite.

Growth rate has been suggested as a factor controlling the inorganic composition of calcified tissues¹¹. A range of growth rates was observed for the mussels cultured in each Mg:Ca ratio solution. Mussels in the two lowest Mg:Ca ratio solutions grew slowly and only a few grew sufficiently to allow sampling (≥ 1 mm new growth). Mussels in the other solutions examined grew at about the same rate (from 1 to 4 mm in length in 8 weeks). Mussels placed in a solution with a Mg:Ca ratio of 13.5 did not grow. There was no correlation between growth rate and the shell Mg:Ca ratio of mussels grown in the same solution. High Mg:Ca ratios did not significantly influence growth rate except in the one case noted above. In this experiment growth rate was not a factor influencing the magnesium composition of *Mytilus* shell calcite.

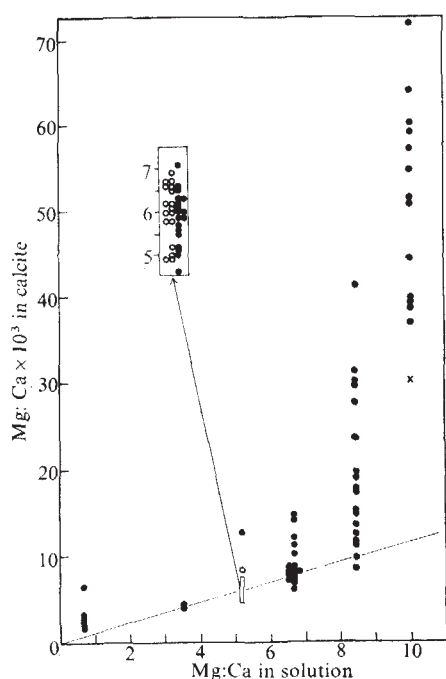


Fig. 1 The calcite shell Mg:Ca atom ratio for each mussel plotted against the Mg:Ca atom ratio of the solution in which it was grown. ○, Samples of mussels grown in natural seawater; ●, mussels grown in synthetic medium. × Represents the metal ratio determined from a composite sample by X-ray diffraction.

Anomalous increases in the magnesium content of these calcite shells could occur if another mineral phase such as brucite ($\text{Mg}(\text{OH})_2$) or hydromagnesite ($3\text{MgCO}_3 \cdot \text{Mg}(\text{OH})_2 \cdot 3\text{H}_2\text{O}$) were precipitated; we carried out X-ray diffraction studies to check this possibility. Mussels grown in 100% natural seawater were almost pure calcite ($d_{211(104)} = 3.035 \text{ \AA}$), whereas a sample taken from mussels grown in the highest Mg:Ca ratio solution had a d spacing of $3.027 \text{ \AA}$. This change in d spacing corresponds to a 2.5 mol per cent increase in MgCO_3 , as calculated from data showing the dependence of d spacing on calcite magnesium content¹². From this increase a Mg:Ca ratio of 31×10^{-3} is calculated for the high magnesium calcite sample. While this ratio falls a little short of the Mg:Ca ratio determined by atomic absorption, it is still several times greater than the ratio that would be obtained by a linear extrapolation from the origin through the normal

seawater ratio (see Fig. 1). The only other peak observed in the high magnesium sample had a d spacing of $17.67 \text{ \AA}$. This d spacing does not correspond to any magnesium containing mineral listed in the tables of the American Society for Testing Materials. High d spacings are often associated with organic material; whether or not this peak represents an elevated shell organic matrix content and whether or not magnesium is associated with it, is conjecture at this point. The X-ray data are in agreement with the atomic absorption data and indicate that an anomalous increase in the coprecipitation of magnesium into *Mytilus* calcite occurs when mussels are grown in solutions containing a higher than normal magnesium content.

Magnesium in high concentrations inhibits the nucleation¹³ and growth⁶ of calcite. Simkiss¹¹ proposes that the mantle of some molluscs may regulate the flow of material to the crystallisation site so as to exclude crystal growth inhibitors such as magnesium and inorganic phosphate ions. We suggest that *Mytilus* does in fact have the capacity to extensively regulate the magnesium ion activity of the solution from which its shell calcite is grown. *Mytilus* may attempt to maintain an optimum magnesium and calcium composition in this fluid. In normal seawater conditions *Mytilus* produces a low magnesium calcite as a result of this ion fractionation process. At very low Mg:Ca solution ratios the ion fractionation process may discriminate less against magnesium so that the normal, or close to the normal, solution composition is maintained. This is suggested by the slightly elevated shell Mg:Ca ratio at the lowest solution ratio. When *Mytilus* is exposed to Mg:Ca solution ratios much above those of normal seawater, its ion regulatory capability apparently becomes overloaded and breaks down, thus allowing an influx of magnesium ions into the fluid from which new shell is precipitated. The result, as shown in Fig. 1, is a dramatic increase in the amount of magnesium found in *Mytilus* shell calcite.

While data are not available on species other than *Mytilus*, it is probable that other marine organisms which produce low-Mg calcite, such as planktonic foraminifera and coccolithophorids, do so by a physiological fractionation process rather than as a result of any inorganic processes.

We thank Pierre E. Biscaye for performing X-ray diffraction work. We also thank the staff of the Rhode Island Nuclear Science Center for their assistance, A. N. Sastry for animal culture facilities, and M. E. Q. Pilson for reviewing the manuscript. N.S.F. provided financial support (DES37125).

ROBERT B. LORENS
MICHAEL L. BENDER

Graduate School of Oceanography,
University of Rhode Island
Kingston, Rhode Island 02881

Received 18 November 1976; accepted 1 September 1977.

1. Milliman, J. D. *Marine Carbonates* (Springer, New York, 1974).
2. Lowenstam, H. A. *J. Geol.* **69**, 241–260 (1961).
3. Thompson, G. & Bowen, V. T. *J. mar. Res.* **27**, 32–38 (1968).
4. Savin, S. & Douglass, R. G. *Geol. Soc. Am. Bull.* **84**, 2327–2342 (1973).
5. Bender, M. L., Lorens, R. B. & Williams, D. F. *Micropaleontology* **21**, 448–459 (1975).
6. Berner, R. A. *Geochim. cosmochim. Acta* **39**, 489–504 (1975).
7. Katz, A. *Geochim. cosmochim. Acta* **37**, 1563–1586 (1973).
8. Kilbourne, R. T. & Sen Gupta, B. K. *S. East. Reg. G. S. A. abstr. programs* 408–409 (1973).
9. Lorens, R. B., Bender, M. L. & Williams, D. F., *J. Sediment Pet.* (in the press).
10. Moller, P. & Parekh, P. P. *Mar. Chem.* **3**, 63–77 (1975).
11. Zolotarev, V. N. *Geochim. Int.* **11**, 347–353 (1975).
12. Goldsmith, J. R. & Graf, D. L. *Am. Mineral.* **43**, 84–101 (1958).
13. Pytkowicz, R. M. *J. Geol.* **73**, 196–199 (1965).
14. Simkiss, K. *Comp. Biochem. Physiol.* **16**, 427–435 (1965).

Polonium and plutonium in an intertidal food chain

LARGE quantities of seaweeds (kelp) are washed ashore along the coasts of California. Estimates range from 10^4 to 10^6 tonnes (10^7 – 10^9 kg) wet weight per yr. These algal masses provide food for insects and other invertebrates on the beaches. The plants carry with them a variety of di-, tri-, and polyvalent cations of heavy metals, largely adsorbed or chelated to superficial acidic polysaccharides and surface slimes of bacterial origin. Among the heavy metals one can detect small amounts of radioactive elements, including some that emit high-energy α particles such as the natural polonium (^{210}Po , half life 138 d) and the artificial plutonium nuclides ($^{239,240}\text{Pu}$, half lives 24,360 yr and 6,580 yr respectively)¹. In surface waters of the oceans, polonium is born from the decay of ^{210}Pb which is continuously replenished by decay of radon released from the Earth's crust. The plutonium isotopes occur in ocean waters as a consequence of fallout from atomic test explosions some 15–20 yr ago. We have examined the extent to which these elements may be transferred from the ocean to the terrestrial biosphere by means of kelp, kelp flies and their predators. Fears have been expressed that radioactivity from the algae could be reaching man through the terrestrial food chain. Our investigations suggest that, on the basis of current radiation thresholds for human health and safety, such fears are groundless.

Samples of cast weed, largely comprising blades and stipes of *Macrocystis pyrifera* (L.) C.A. Ag. but containing also a few per cent of *Egria laevigata* Setch. and *Pelagophycus porra* (Leman) Setch. and marine grasses (*Phyllospadix* and *Zostera*), were collected from the beach at La Jolla, California, in June and August 1975, at a time when kelp flies were much in evidence. Adult flies (mostly *Fucilia rufitibia* Stein and *F. separata* Stein) were caught over clumps of decaying weed by the use of a 1-mm mesh nylon net. About 30,000 flies, weighing 100 g wet weight (27 g dry weight) were collected in about 30 min. A few of the associated predatory tiger beetles (*Cicindela* sp.) were also netted, as individuals, on nearby sand. For comparative purposes we also obtained some 50 g (dry weight) of common houseflies (*Musca domestica* L.) which had been raised in a laboratory (Department of Entomology, University of California, Riverside) on a commercial fly-larva medium (CSMA, Ralston-Purina Co.).

All samples were dried in an oven at 80 °C overnight. Polonium and plutonium were determined in subsamples of 1.0 g and 20.0 g respectively. The flies and beetles were 'wet-ashed' by heating in a mixture of nitric and perchloric acids. As markers, ^{208}Po and ^{242}Pu were added to monitor the chemical recoveries of ^{210}Po and $^{239,240}\text{Pu}$ from the samples. The α activities of these elements were measured with silicon surface-barrier detectors in combination with pulse-height analysers.

As shown in Table 1, the contents of polonium in laboratory-raised houseflies, and in the diet on which they were raised as larvae, were negligible (10 and 30 pCi per kg dry weight) compared with the kelp flies (250 pCi per kg dry weight). The polonium content of the kelp flies from the beach was found to be about half of that of the cast weed, on the basis of comparable dry weights. (That of laboratory-reared *Musca* was about one third that of their larval food.) This proportional reduction may be attributable to mechanical selection by the fly larvae of the softer, interior parts of the seaweeds, which have been shown to contain considerably less of the adsorbed heavy metals (including polonium) than the surface layers². Alternatively, there may have been discrimination against uptake of polyvalent metallic cations in the insect gut, selective metabolic processes in the tissues, or active excretion of such cations.

The polonium content in the predatory tiger beetles was found to be about one-half that of kelp flies, a fact which may be attributable to similar mechanical or metabolic discrimination. (It would be interesting to determine the polonium content of other predators of kelp flies, for example birds such as spotted sandpipers (*Petrochelidon pyrrhonota*) or swallows (*Hirundo rustica*), but we did not attempt this.) In neither of the two successive steps

Table 1 Polonium and plutonium contents of algae and flies from a California shore

		^{210}Po	$^{239,240}\text{Pu}$	Po/Pu	^{210}Pb
Beach					
Kelp:	<i>Macrocystis</i>	480 $\pm$ 40	2.0 $\pm$ 0.3	250	380 $\pm$ 40
	and others	510 $\pm$ 25			
Flies:	<i>Fucilia</i> : whole adults	290 $\pm$ 15	1.1 $\pm$ 0.2	260	220 $\pm$ 30
	hard parts	247 $\pm$ 18			
	expressed 'juice'	277 $\pm$ 13			
		< 5			
Beetles:	<i>Cicindela</i>	164 $\pm$ 12			
Laboratory					
Fly-larva food (CSMA)		30 $\pm$ 3			
Flies:	<i>Musca</i> : whole adults	10 $\pm$ 4			

Values are pCi per kg dry weight; shown as mean $\pm$ s.e.m.

of the food chain studied here did we observe a relative enrichment of the radionuclides. On the contrary, at each step the relative polonium concentration was halved, while the plutonium (and ^{210}Pb) content of the kelp flies was also only half that of the content of the decayed kelp on which they fed.

A crude extract of the blood, gut contents and other soft parts of the kelp flies was prepared by homogenising a sample of the fresh flies in a Waring blender with distilled water for 45 s and centrifuging at 13,000 r.p.m. (20,000g) for 10 min. The liquid fraction was then clarified by filtration. The residues retained by the filter and at the bottom of the centrifuge tube consisted of cuticles, heads, wings, legs and bristles. This fraction contained essentially all of the polonium of the intact flies (see Table 1), indicating that most of the polonium ingested by the insect larvae with their food (the decaying kelp) had been incorporated in the cuticles of the adult flies.

In ascending the food chain from the kelp to the flies, the Po/Pu ratio remained unchanged. In contrast, corresponding values for *Macrocystis* plants and for bryozoa removed from their surfaces were respectively 230 and 1,000¹, while even larger ratios, as high as 10^5 or 10^6 , were found in the internal organs of several species of marine fishes (V.F.H., unpublished), suggesting that in these latter systems there is a more effective discrimination against Pu, or in favour of uptake of Po, or both, than in the kelp flies.

Our observations indicated that on a warm summer day there may be as many as 10^7 flies per km of beach near the Scripps Institution of Oceanography, in an area where the estimated annual tonnage of cast weed is about 10^5 kg km⁻¹ (ref. 3). The kelp brings with it some 10^7 pCi of polonium and 4×10^4 pCi of plutonium per km of beach annually. If we assume that there are 10 generations of kelp flies per yr (in summer⁴ a generation time of kelp flies may be as short as 10 d), it can be calculated that about 0.2% of the polonium and 0.2% of the plutonium adsorbed or in some other way fixed to the major seaweeds may find its way into a terrestrial food chain.

It should be emphasised that although the kelp flies may occur in large numbers along beaches on warm days, such as those observed during the great seaweed-fly plague in southern England in 1953–54⁵, the radiation values with which we are concerned here present no obvious hazards to human health. The accidental inhalation of a few kelp flies would involve the ingestion of merely 10^{-4} pCi of ^{210}Po and 10^{-6} pCi of $^{239,240}\text{Pu}$, while the deliberate ingestion of tens of thousands of such flies would be too involved and unappetising an exercise to consider seriously.

We thank Dr Willis W. Wirth (Smithsonian Institution, Washington) for identifying the kelp flies, and Mr Ian Moore (University of California, Riverside) for providing laboratory-reared houseflies and the larval medium.

LANNA CHENG
V. F. HODGE
R. A. LEWIN

Scripps Institution of Oceanography,
University of California,
La Jolla, California 92093

Received 11 July; accepted 15 August 1977.

1. Hodge, V. F., Hoffman, F. L. & Folsom, T. R. *Health Phys.* **27**, 29 (1974).
2. Wong, K. M., Hodge, V. F. & Folsom, T. R. *Nature* **237**, 460 (1972).
3. Zobel, C. E. *Nova Hedwigia* **32**, 269 (1971).
4. Dobson, T. in *Marine Insects* (ed. Cheng, L.) 447 (North-Holland, Amsterdam, 1976).
5. Oldroyd, H. *Discovery* **15**, 198 (1954).

Free-running activity rhythm in the natural environment

RHYTHMICAL activities of plants and animals in their natural habitats have, so far, always been found to be synchronised with cyclic environmental changes¹⁻³. In the absence of appropriate environmental timing cues (usually light-dark or tidal cycles) the rhythms free-run with endogenously controlled periods other than those of the environmental cycle. These may be longer or shorter than those observed in the presence of the entraining environmental cycles. We describe here an apparently unique example of an invertebrate which seems to free-run in its natural habitat. The animal is a predatory prostigmatid mite (*Bdella interrupta* Evans) which inhabits marine salt marshes and is exposed to regular tidal inundation during periods of spring tides. During periods of neap tides the salt marsh is not covered by the sea. The mites showed day and night peaks of locomotory activity on the soil surface at creek edges (Fig. 1a) and, also, in isolated experimental conditions (Fig. 1b). The narrow peaks of activity imply that the individual members of the population are closely synchronised with one another.

The peaks of locomotory activity in the natural habitat did not exhibit constant relationships to the times of dawn and dusk. The interval between day peaks of activity during a sequence of non-covering tides was less than 24 h. The slope of the regression line for the first 6 d illustrated in Fig. 2 corresponds to an average separation of 23.1 h between successive day peaks during this period of non-submerging tides. During

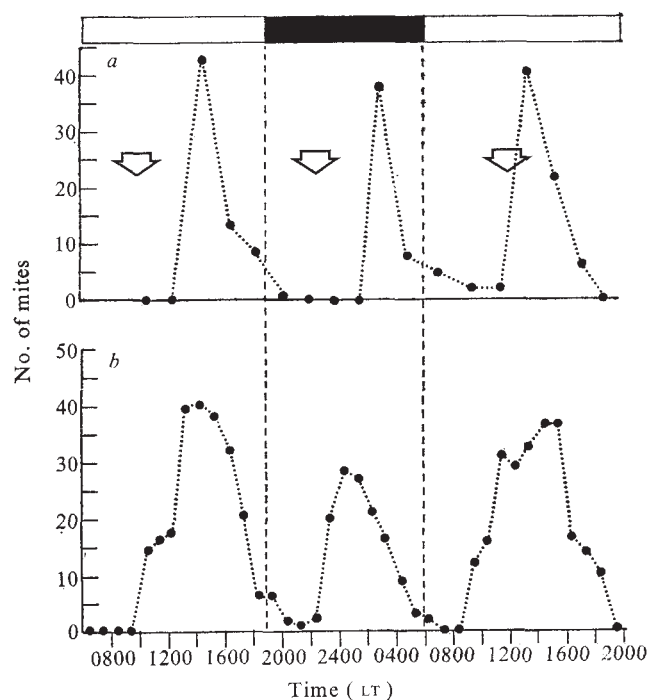


Fig. 1 *a*, Numbers of mites observed walking on the soil surface on the edges of a drainage channel on Hut Marsh, Scolt Head, Norfolk, UK. Observations were made on a 25-m length of bank during a period of tidal emergence (6-7 April 1976). The times of high tides are indicated by open arrows. Dawn and dusk are shown by broken lines. *b*, Surface activity of *B. interrupta* in isolated habitat portions (in five cylindrical containers of 15.2 cm diameter containing soil of 10 cm depth). The mites were maintained in natural light conditions.

the subsequent period of submerging tides the average separation between successive day peaks increased to 25.1 h, so that the surface locomotory activity of the mites effectively paralleled the times of the high tides for the period (Fig. 2).

Periodograms for the 1976 data illustrated in Fig. 2 yielded activity peaks at 22.9 h (for locomotory activity during non-submerging tides) and 24.9 h (during the subsequent period of submerging tides) (Fig. 3) for data collected during the daylight hours. Field observations made in 1977 confirmed these differences in periodicity observed during periods of tidal emergence and submergence. Since the activity peaks occur twice

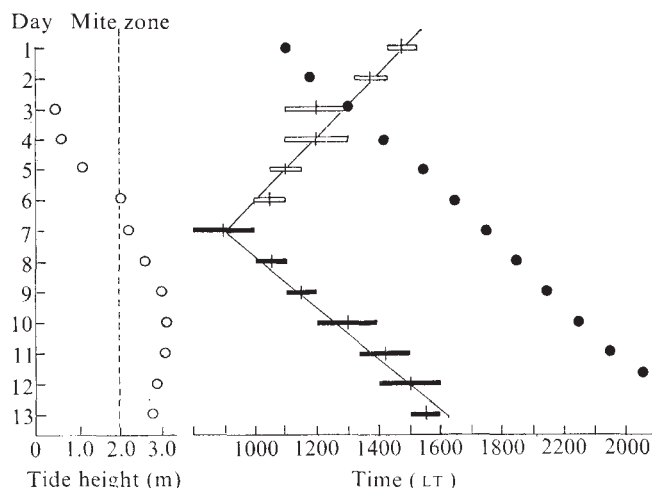


Fig. 2 The timing of day activity peaks related to the time (●) and the height (○) of high tides from 6 to 18 April 1976. The horizontal bars indicate the period during which mite activity was more than 50% of the maximum. The vertical lines are the midpoints of the bars and are taken as the phase reference point. The open bars represent activity during tidal emergence and the closed bars that during a subsequent period of submerging spring tides. The continuous lines are the calculated regression lines (for the period of tidal emergence: $r = 0.9720$; $n = 6$; $P < 0.005$; for the period of tidal submergence $r = 0.9923$; $n = 7$; $P < 0.001$). For these calculations the day sequence was used as the x axis.

daily and the rhythm entrains to the tidal frequency when the animal is exposed to tidal inundation the locomotory activity rhythm can be defined as a circatidal rhythm with a periodicity of about 12.5 h. As with circatidal rhythms in other organisms, the mite rhythm is insensitive to light-dark cycles. In this respect this salt-marsh mite is similar to an oribatid mite (*Ameronothrus marinus* Banks), which on rocky shores shows an apparently circatidal rhythm of 12.3-h periodicity⁴.

The novel feature revealed by this investigation is that during periods of non-submerging tides the period of the activity rhythm (11.5 h) is shorter than when it is entrained by the tides (12.5 h) and can be reasonably presumed to be a free-running circatidal rhythm. The maintenance of narrow peaks of activity throughout periods of tidal emergence indicates that the periods of the individuals are very similar whilst free-running. This contrasts with the breakdown of synchrony seen in free-running populations of other organisms in constant experimental conditions.

There is no conclusive evidence that the alternation of an apparently free-running rhythm, during tidal emergence, with an entrained rhythm, during periods of tidal submergence, is of adaptive value. It is conceivable, however, that this alternation could be a mechanism to achieve synchrony with the pattern of tidal coverage on the marine salt marsh. At the onset of periods of tidal coverage the activity peaks occur approximately midway between the times of successive high tides. This temporal separation is maintained during subsequent periods of tidal coverage by the circatidal rhythm of 12.5-h periodicity. The alternation between 11.5- and 12.5-h periodicities ensures

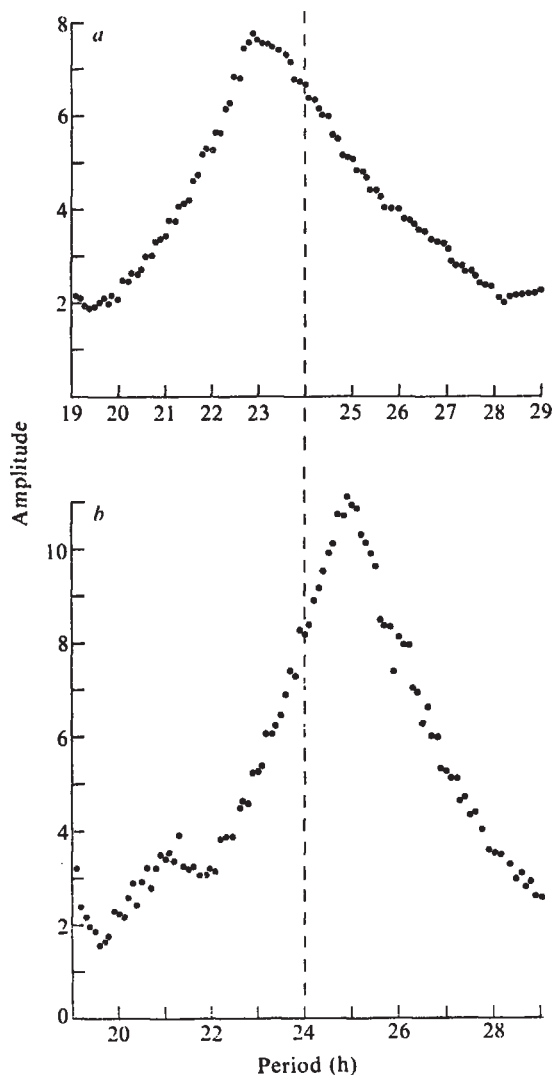


Fig. 3 Periodograms of data illustrated in Fig. 2, for the observed locomotory activity during periods of non-submerging (a) and submerging (b) tides.

that a peak of mite activity is continuously maintained in daylight hours. This would not occur if the circatidal periodicity (12.5 h) continued during tidal emergence. Calculations using data from the 1977 tide tables show that with an intermediate free-running period, for example 12.0 h, the activity peak is always within 2.5 h for the first, critical, tidal inundation in a series of rising spring tides. With a free-running period of 11.5 h, on the other hand, the activity peak is within 2.5 h of only 17% of the critical tides.

We thank Drs Jack Benson and J. Brady for helpful discussions. We also thank the Nature Conservancy for permission to work on the Scolt Head Nature Reserve.

J. E. TREHERNE
W. A. FOSTER
P. D. EVANS
C. N. E. RUSCOE*

Department of Zoology,
University of Cambridge, UK

Received 25 July; accepted 19 August 1977.

*Permanent address: ICI Plant Protection Division, Jealott's Hill Research Station, Bracknell, Berkshire, UK.

1. Bünning, E. *The Physiological Clock* (English Universities Press, London, 1973).
2. Enright, J. T. *J. comp. Physiol.* 77, 141-162 (1972).
3. Palmer, J. D. *Biol. Rev.* 48, 377-418 (1973).
4. Schulte, G. *Mar. Biol.* 37, 265-277 (1976).

Sexual dimorphism, socionomic sex ratio and body weight in primates

BECAUSE the primates are a particularly well studied group they provide a rare opportunity to investigate the adaptive significance of species differences in sexual dimorphism in body size. We describe here an investigation of the relationship between the degree of sexual dimorphism and three variables which are predicted might affect it.

In many birds and mammals (with certain exceptions)^{1,2} the male is larger than the female. The commonest explanation for this trend, originally offered by Darwin in 1871 (ref. 3) is that increased body size confers advantages in contests between males for access to females (and there is now detailed evidence supporting this⁴⁻⁶). The basis of this argument is that males are able to enhance their breeding success greatly by maintaining exclusive breeding access to several mates³, whereas females cannot usually⁷ enhance theirs to the same extent by maintaining exclusive access to multiple males. For this reason, the argument runs, polyandry is rarer than polygyny. Once polygyny has developed, males will compete for access to females⁴ and individuals possessing traits which enhance their competitive ability (such as increased body size) will be selected. This argument assumes that increases in body size are constrained by ecological, perhaps energetic, disadvantages which can be outweighed by reproductive advantages.

There is another functional explanation of sex differences in body size—these may permit the sexes to utilise different kinds of food and consequently enhance the efficiency with which they can exploit their resources. These two theories are not mutually exclusive, but their predictions differ: the first predicts that sexual dimorphism should be most marked in strongly polygynous species where competition between males for access to females is intense since only a small proportion of males can breed; the second, that dimorphism should be most marked in monogamous species or in those with a relatively equal adult sex ratio where divergence between the sexes would be most effective^{8,9}. While there is reasonable evidence implicating the second explanation in particular cases⁸, the distribution of dimorphism across species generally fits the predictions of the first: monogamous species show least dimorphism and those with the most imbalanced socionomic sex ratios (the number of adult females per adult male in breeding groups) the most.

But, although a positive relationship between body size dimorphism and socionomic sex ratio has been suggested in reviews of birds^{1,10,11}, ungulates^{12,13} and primates¹⁴⁻¹⁶, none have offered formal statistical evidence for such a relationship. That such evidence is necessary is demonstrated by the extent to which species with similar breeding systems differ in their degree of body size dimorphism (see below). Also, it is necessary to investigate what other factors affect the degree of size dimorphism and to consider the possibility that the apparent relationship between dimorphism and socionomic sex ratio is the product of some third factor associated with these two variables.

In most orders, formal comparisons are impossible because of the paucity of accurate data about social organisation. This is not so in the primates, where the social behaviour of a large number of species has now been studied in detail. To examine this relationship, we extracted as many estimates of body weight and of breeding sex ratio from the literature as we could find (the sources and data used will be published elsewhere (refs 17, 18 and B. Rudder, in preparation), selecting weights of wild animals in preference to those collected from zoos wherever possible. Sexual dimorphism was then calculated by dividing male weight by female weight. Our measure of socionomic sex ratio was the mean number of adult females per adult male in breeding groups: where several estimates were available for the same species, we took the mean of these. Our sample included 42 primate species.

There is a significant relationship between sexual dimorphism and socionomic sex ratio (Fig. 1: $r_s = 0.610$, $t = 4.87$, 40 d.f., $P < 0.001$) but, this relationship is largely due to the fact that

monogamous species show consistently less dimorphism than polygynous ones (Kruskal-Wallis $H = 22.16$, $P < 0.001$) and, when monogamous species are excluded, no relationship exists ($r_s = 0.087$, $t = 0.44$, 27 d.f., not significant). Why should this be the case? There are probably two main reasons. First, socionomic sex ratio fails to provide a reliable index of inter-male competition. It would, for example, be likely to underestimate the degree of competition in multi-male troops (as in *Papio* or *Macaca* species) where a marked dominance hierarchy exists and breeding access is not shared equally among the males. Conversely, it would overestimate the degree of competition in species where access to females is strongly age-dependent and the majority of males pass through a short reproductive phase¹⁹. Also, our data will inevitably have included a proportion of cases where estimates of socionomic sex ratio were a typical, thus introducing error. Second, the data show that factors apart from socionomic sex ratio exert a strong influence on the degree of sexual dimorphism. Terrestrial species show consistently greater dimorphism than arboreal ones (Kruskal-Wallis $H = 8.96$, $P < 0.01$) probably because, in the former, increased weight does not restrict the male's access to food supplies (often located on fine terminal twigs) to the same extent as in the latter. In addition, sexual dimorphism is positively related to body weight (the mean of adult male and adult female body weight: see Fig. 2: $r_s = 0.484$, $t = 3.50$, 40 d.f., $P < 0.01$). This relationship is not a product of the inclusion of female weight in both parameters: if male weight is plotted against

female weight, the slope is significantly greater than 1 ($b = 1.063$, $t = 66.33$, $P < 0.001$). Nor is it the product of either of an association between large-size and terrestriality since the same relationship holds within arboreal species (Fig. 2: $r_s = 0.395$, $t = 2.27$, 28 d.f., $P < 0.05$) or of an association between size and polygyny (see below).

Since several monogamous species are of small size (such as *Callicebus* spp., *Saguinus* spp.) it is necessary to consider whether the overall relationship between dimorphism and socionomic sex ratio could be the product of that between the former and body weight. To do this, we linearised the relationships and used a partial regression analysis²⁰. This tests whether a dependent variable (in this case, sexual dimorphism) is significantly related to each of two or more independent variables (socionomic sex ratio and body weight) when the effect of the other has been removed. (In the description of results, b_1 refers to the regression coefficient for the first independent variable and b_2 to the second). The analysis showed that sexual dimorphism is significantly related to both variables in the sample as a whole ($b_1 = 0.111$, $t = 2.14$, $P < 0.05$; $b_2 = 0.078$, $t = 3.11$, $P < 0.005$, 39 d.f.) and that 56.3% of the variance in sexual dimorphism is accounted for by the two variables. The relationship to body weight was considerably stronger than that to socionomic sex ratio: standardised partial regression coefficients for sex ratio and body weight are 0.28 and 0.41 respectively. Similar results were obtained when we applied the same analysis to the data for arboreal primates only

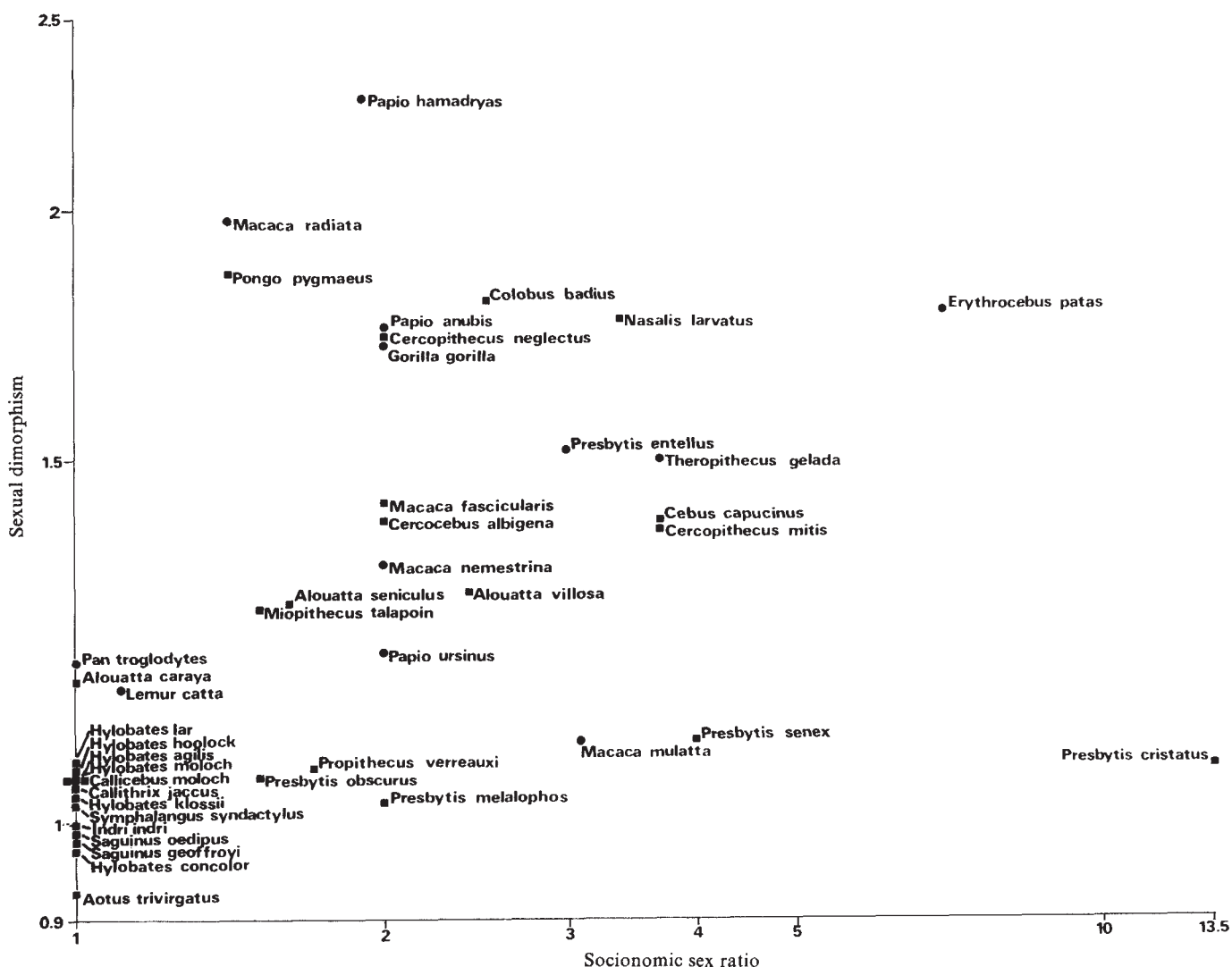


Fig. 1 Sexual dimorphism in body weight (weight of male/weight of female) plotted against socionomic sex ratio (females per male). ■, Arboreal; ●, terrestrial.

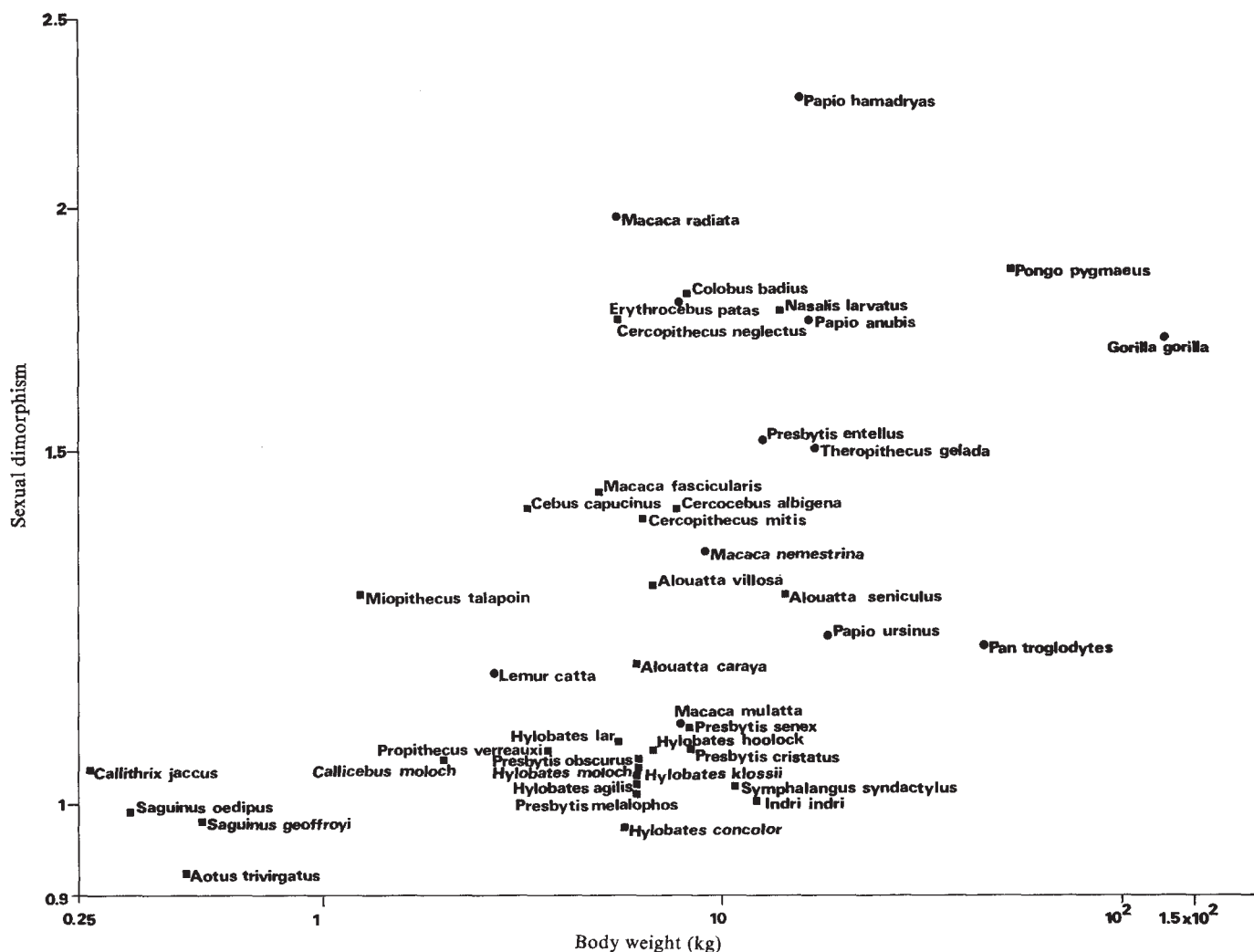


Fig. 2 Sexual dimorphism in body weight plotted against average of adult male and adult female body weight. ■, Arboreal; ●, terrestrial.

($b_1 = 0.097$, $t = 1.81$, $P < 0.1$; $b_2 = 0.064$, $t = 2.32$, $P < 0.05$, 27 d.f.).

Why should sexual dimorphism in body weight be positively correlated with absolute body weight? There is extensive evidence that this relationship is not confined to primates and occurs in invertebrates²¹, birds^{10,21,22}, rodents (G. Mace, personal communication) and carnivores. It was apparently first recognised by Rensch^{21,23} who attributed it to allometric growth: "Secondary sexual characters of many mammals usually represent structures of positively allometric growth because they usually become more conspicuous during a more advanced period of ontogenetic development (due in mammals to the increasing influence of sex hormones)".

There is, however, no obvious reason for supposing that allometry is inevitable. At least six functional explanations could be involved. It has been suggested that sexual selection for increased size in males may, as a side effect, have led to increased female size thus creating a relationship between dimorphism and average body size²⁴. This seems unlikely because it requires the assumption that increases in male size are more closely paralleled by increases in female size in small animals than in large. Moreover, selection for energetic efficiency probably imposes strict constraints on female size^{9,10,18,25}. Alternatively, it is generally true that large body size is associated with a reduction in the number of competing species^{26,27}. This could relax constraints on increasing male body size imposed by the presence of larger species in neighbouring niches, allowing sexual selection for male competitive ability to produce relatively large increases in male

size. There is evidence that in some birds sexual dimorphism in body size is commonly associated with sex differences in food choice (see above) and that the presence of competing species can constrain intraspecific variance in many characteristics, including size²⁶⁻²⁸. It is also possible that energetic constraints on increasing male body size may be relaxed in larger species since larger body size leads to a decrease in the surface area: volume ratio, metabolic rate and relative nutritional requirements (refs 30-32 and J. Kurland, personal communication) thus making further increases energetically 'cheaper'. Alternatively, where dimorphism permits the sexes to diverge in food choice, selection may favour increased feeding divergence (and thus increased dimorphism) in species which feed on low density food distributed in clumps²² since, in this situation, the effects of feeding competition are strongest. Mammalian species which feed on dispersed and unpredictable food supplies are often of large size^{12,32}. Finally, it is possible that evolutionary factors favouring the reduction of female size may be involved³³. It has been argued that, in fluctuating environments, smaller females may be able to breed at a younger age than larger ones³⁴ and it is conceivable that this pressure might be more acute in large species where the age at first breeding is relatively late, than in small species where it is generally early.

None of these explanations is wholly satisfactory and evidence to support them is almost totally lacking. Until they have been investigated systematically, however, it would be premature to assume that the relationship between body weight and sexual dimorphism is inevitable.

We thank Dr S. A. Altmann, Dr A. H. Harcourt, R. M.

Gibson, P. Greene, P. J. Greenwood, G. Mace, Professor J. Maynard Smith, Dr N. Owen-Smith and Professor R. L. Trivers for help, advice and criticism.

T. H. CLUTTON-BROCK

King's College Research Centre, Cambridge, UK

PAUL H. HARVEY

School of Biology, University of Sussex, Brighton, UK

B. RUDDER

Wellcome Institute of Comparative Physiology, Regent's Park, London, UK

Received 25 January; accepted 13 August 1977.

1. Lack, D. *Ecological Adaptations for Breeding in Birds* (Methuen, London, 1968).
2. Ralls, K. *Q. Rev. Biol.* **51**, 245–275 (1977).
3. Darwin, C. *The Descent of Man and Selection in Relation to Sex* (Murray, London, 1871).
4. Trivers, R. L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B.) (Aldine, Chicago, 1972).
5. Wilson, E. O. *Sociobiology, the New Synthesis* (Belknap, Cambridge, Massachusetts, 1975).
6. Geist, V. *Mountain Sheep* (University of Chicago Press, Chicago, 1971).
7. Jenni, D. A. *Am. Zool.* **14**, 129–144 (1974).
8. Selander, R. W. in *Sexual Selection and the Descent of Man* (ed. Campbell, B.) (Aldine, Chicago, 1972).
9. Bourlière, F. in *Small Mammals, their Productivity and Population Dynamics* (eds Golley, F. B., Petrusewicz, K. & Ryszkowski, L.) (Cambridge University Press, Cambridge, 1975).
10. Wiley, R. H. *Q. Rev. Biol.* **49**, 201–227 (1974).
11. Crook, J. H. *Behav. Suppl.* **10**, 1–178 (1964).
12. Jarman, P. J. *Behaviour* **48**, 215–267 (1974).
13. Geist, V. *Am. Zool.* **14**, 205–220 (1974).
14. Crook, J. H. in *Sexual Selection and the Descent of Man* (ed. Campbell, B.) (Aldine, Chicago, 1972).
15. Gautier-Hion, A. *Mammalia* **39**, 365–374 (1975).
16. Clutton-Brock, T. H. & Harvey, P. H. in *Growing Points in Ethology* (eds Bateson, P. P. G. & Hinde, R. A.) (Cambridge University Press, Cambridge, 1976).
17. Clutton-Brock, T. H. & Harvey, P. H. *J. Zool.* **183**, 1–39 (1977).
18. Clutton-Brock, T. H. & Harvey, P. H. in *Primate Ecology: Studies of Feeding and Ranging Behaviour in Lemurs, Monkeys and Apes* (ed. Clutton-Brock, T. H.) (Academic, London, 1977).
19. Rowell, T. E. *Behav. Biol.* **11**, 131–154 (1974).
20. Bailey, N. T. J. *Statistical Methods in Biology* (English Universities Press, London, 1964).
21. Rensch, B. *Evolution Above the Species Level* (Methuen, London, 1959).
22. Selander, R. K. *Candor* **68**, 113–151.
23. Rensch, B. *Bonner Zool. Beitr.* **1**, 58–69.
24. Maynard Smith, J. *Anim. Behav.* **25**, 1–9.
25. Bell, R. H. V. *Sci. Am.* **225**, 86–93 (1977).
26. MacArthur, R. H. *Geographical Ecology* (Harper and Row, New York, 1972).
27. MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton University Press, Princeton, 1967).
28. Yom Tov, Y. & Ollason, T. G. *Oikos* **27**, 81–85 (1976).
29. Lack, D. *Ecological Isolation in Birds* (Blackwells, Oxford, 1970).
30. Kleiber, M. *The Fire of Life* (Wiley, New York, 1961).
31. Janis, C. *Evolution* **30**, 757–774 (1976).
32. Clutton-Brock, T. H. in *Primate Ecology: Studies of Feeding and Ranging Behaviour in Lemurs, Monkeys and Apes* (ed. Clutton-Brock, T. H.) (Academic, London, 1977).
33. Ralls, K. *Am. Nat.* (in the press).
34. Downhower, J. F. *Nature* **263**, 558–563 (1976).

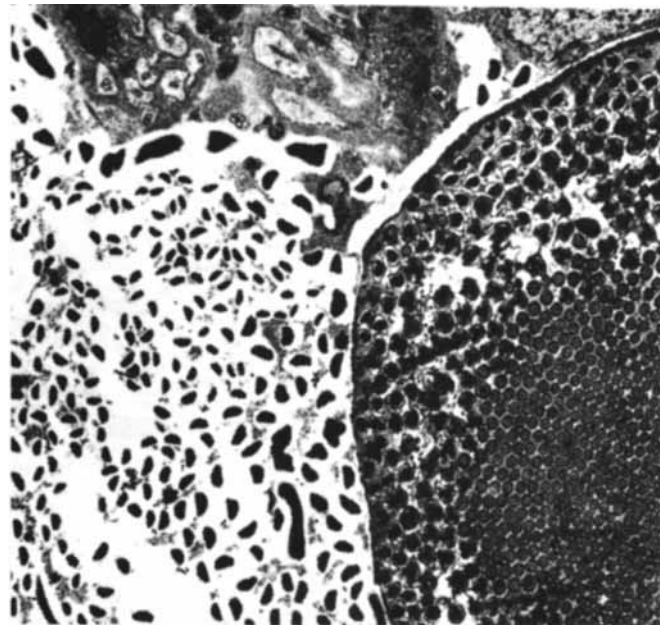


Fig. 1 Transverse section of sperm duct. Spermatophore containing mature spermatozoa cut at various levels (left), electron-dense threads (right) and parts of the filament (bottom). Phosphate-buffered glutaraldehyde (3.5%) and osmium (2%) fixation. Araldite embedding, uranyl acetate and lead citrate staining ($\times 6,080$).

forceps. Mature males were identified easily by the large numbers of spermatophores within the transparent sperm ducts.

Sections cut at various levels of the sperm ducts show the well known spindle-shaped spermatophores, portions of their coiled filaments and numerous electron-dense threads of unknown function (Fig. 1). These threads are about 0.5 μm in diameter and several μm long, their irregular shape preventing exact measurement. High magnifications reveal a new structural element, which I call 'dense core tubuli'. These tubuli constitute the major structural unit of the spermatophore wall as well as the long filament and the electron-dense threads (Fig. 2). The tubuli are embedded in an amorphous ground substance. They have an outside diameter of 150–170 $\text{\AA}$, thus being slightly smaller than the microtubules of many plant and animal cells. In contrast to microtubules, they are clearly extracellular. Their length is indeterminate but tubules of several μm have been

Ultrastructure of the spermatophores of *Siboglinum ekmani* Jägersten (Pogonophora)

THE marine invertebrate phylum Pogonophora has attracted much attention since its discovery in 1914 because of its uncertain phylogenetic position and its unusual biological features, such as the lack of a digestive tract^{1,2}. Little is known of the spermatophores and the mechanism of fertilisation^{1,3}, although it has been suggested that the long filament of the spermatophore serves as a flotation device⁴. Because the expulsion of spermatophores into the sea would involve great wastage of sperm, however, some form of copulation has been said to be likely⁵. Expulsion of spermatophores by the tentacles of the male has been observed in one species³, but even in dense populations of 200 animals per m^2 , which have been reported¹, the tentacles are too short to transfer the spermatophores to the females. Electron microscopy has now revealed that the spermatophoral filament is equipped with numerous electron dense 'threads', composed of tubular elements with a central electron dense core. Apparently the 'threads' which receive their strength from the tubular elements, play an important part as a floating or adhesive apparatus of the released spermatophores.

Several specimens of *Siboglinum ekmani* Jägersten were collected from the Skagerrak during tests of the German research vessel Poseidon (January 1977, 57°50'N, 8°05'E, 537 m deep, 3.4°C, salinity 32.6 ‰). The animals were removed carefully from their tubes using pointed watchmaker's

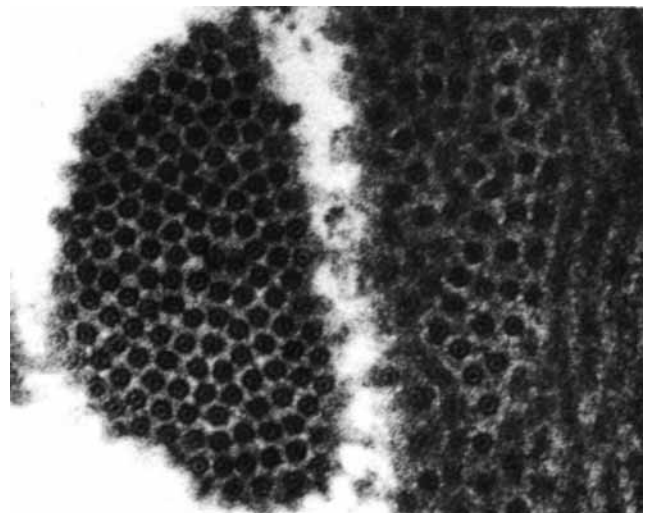


Fig. 2 Cross sections of electron dense thread (left) and spermatophore wall (right) reveal the dense core tubuli ($\times 152,000$).

observed. Occasionally the central core is composed of sub-units which also form a tubular structure. Dense core tubuli are linked in regularly spaced arrangements: 12–180 tubules are longitudinally orientated within the threads. In the spermatophore wall the arrangement resembles dense connective tissue of vertebrates.

Clusters of dense core tubuli also occur between developing male germ cells in the testes. Regularly linked tubuli are attached to groups of mature spermatozoa to mould the envelope of the spermatophores in the proximal region of the sperm ducts. Released spermatophores tend to stick firmly to the tube or trunk of the animal even after fixation, dehydration and embedding. This indicates remarkable adhesive properties. When spermatophores are exposed to seawater, the regular arrangement of the dense core tubuli is gradually lost and the spermatophore wall starts to disintegrate⁶. The filament



Fig. 3 Two stalked vesicles of the spermatophoral filament. Delicate spines project from the surface into the fibrous cover ($\times 30,500$).

uncoils, and at intervals it has areas of radiating stalked vesicles (Fig. 3). Within the sperm ducts these vesicles are smaller, from which I conclude that the filament uncoils through the swelling of the vesicles. They are probably identical with the capitate processes revealed by light microscopy^{3,7}. Fibrous, sticky material covers the surface of the vesicles, reinforced by delicate spines (Fig. 3). The dense threads found in the sperm ducts are also released and frequently appear near the spermatophores in sections. Undoubtedly, the stalked vesicles of the filament and the dense threads form an effective adhesive apparatus and floating device.

My investigation adds further support to the view that the released spermatophores of *Siboglinum* find their way passively without 'some form of copulation' to a receptive female^{1,5}. But further observations are required to elucidate the path of the released spermatozoa to the eggs.

H. FLÜGEL

Institut für Meereskunde,
Universität Kiel, 23 Kiel, FRG

Received 11 July; accepted 23 August 1977.

1. Southward, E. C. *Oceanog. mar. Biol. A. Rev.* 9, 193–220 (1971); in *Reproduction of Marine Invertebrates* (eds. Giese, A. C. & Pearse, J. S.) (Academic, New York, London, 1975); *Symp. Zool. Soc., Lond.* 36, 235–251 (1975).
2. Franzén, A. *Acta Zool.* 54, 179–192 (1973).
3. Webb, M. *Sarsia* 13, 45–49 (1963).
4. Ivanov, A. V. *Pogonophora* (Academic, New York, London, 1963).
5. Carlisle, D. B. in *Pogonophora* (ed. Ivanov, A. V.) (Academic, New York, London, 1963).
6. Southward, A. J. & Southward, E. C. *J. mar. biol. Ass. U.K.* 43, 57–64 (1963).
7. Ivanov, A. V. *Zool. Jb. Abt. syst.* 85, 430–500 (1957).

Blood-sucking flies and primate polyspecific associations

PRIMATE polyspecific associations are likely to be an effective means of reducing the number of fly bites individual monkeys receive¹. The probability of acquiring a vector-borne disease, and the severity of that disease increases with the number of fly bites an animal receives^{2–4}. Bait animals placed in groups are known to attract fewer individual mosquitoes per bait individual than are bait animals placed singly^{3,5}. Mean group size of the mangabey, *Cercocebus albigena* is about 13 in Kibale Forest, Uganda, and the mean group size of *Colobus badius* and *Cercopithecus ascanius* are 50 and 35 respectively⁶. Associations of groups of these species are likely to reduce significantly the number of dipteran bites individual monkeys receive^{1,7}. Other explanations proposed for primate polyspecific associations include increased food location abilities, and increased predator detection and avoidance^{8,9}. I report here the correlation of the temporal occurrence of *C. albigena* polyspecific associations with the activity of biting and sucking flies, *C. albigena* feeding and other activities, and the temporal occurrence of attacks by the crowned eagle-hawk (*Stephanoaetus coronatus*) (eagles are the only known predator of these monkeys). The study was carried out in the Kibale Forest, Ngogo Reserve.

On 5 d each month from January to September 1975, I followed a group of *C. albigena* from dawn to dark. Data on mangabey activity and polyspecific associations were recorded during four 5-min periods each hour (0–5, 15–20, 30–35 and 45–50 min). During each 5-min period as many individual *C. albigena* as possible were located, and the predominant activity of each during a 10-s interval was recorded. 'Moving', 'sitting', and 'feeding' are the activity categories used here. The location of each *C. albigena* observed was plotted on a map (4 cm = 100 m) of the study area. To record daily and hourly movements, net group displacements were measured from one 5-min period to the next. The actual distance measured was from the central focus of individuals in one 5-min period, to the central focus of the next. A polyspecific association was defined as the presence of more than one individual of another species within 20 m of the nearest mangabey. Times of eagle attack were recorded as they occurred. Attacks took the form of an eagle flying fast and low through the tree tops and branches, usually through the midst of the mangabey group. Activity of biting and sucking flies was determined using myself as bait. On 4 d I placed myself (naked arms and legs) on a platform 20 m above the forest floor, and remained there from dawn to dark. All flies (mosquitoes and others) biting me were recorded each hour, and when possible I killed all flies actually biting. Data from all days have been aggregated.

The *C. albigena* group took part in polyspecific associations on 43 of the 45 d of observation. Of a total 513 associations, 90.8% involved *C. ascanius*, 22.01% *C. badius*, 5.07% *Colobus guereza*, 1.75% *Papio anubis* and 0.59% *Cercopithecus thoeisti*. There was little interaction between species during the associations. On two occasions juvenile *C. albigena* attempted to join play groups of the other species. Other interactions were confined to adult *C. albigena*, displacing or occasionally chasing *C. ascanius*.

The occurrence of *C. albigena* in polyspecific associations was not equally distributed throughout the day ($\chi^2 = 19.71$, d.f. = 9, $P < 0.05$) (Table 1). There was a significant correlation between the hourly occurrence of polyspecific associations and the hourly activity of biting flies ($r_s = 0.63$, $N = 10$, $P < 0.05$) (Table 1). Because the data on Kibale fly activity are sparse, I also compared the temporal pattern of associations with the timing of mosquito activity in rainforest in Bwamba County, Uganda¹⁰. The correlation is significant ($r_s = 0.66$, $N = 10$, $P < 0.05$) (Fig. 1).

There were no significant correlations between the hourly occurrence of primate polyspecific associations and eagle attacks ($r_s = 0.119$, $N = 10$, $P > 0.05$), mangabey feeding activity ($r_s = 0.08$, $N = 10$, $P > 0.05$), sitting activity ($r_s = -0.08$, $N = 10$, $P > 0.05$), movement activity ($r_s = 0.42$, $N = 10$, $P > 0.05$), and the hourly distance travelled ($r_s = -0.05$, $N = 10$, $P > 0.05$) (Table 1).

Table 1 Polyspecific associations, feeding moving and sitting activities, hourly distances travelled, eagle attacks and the number of biting flies of a group of mangabeys

Time	Total associations	Feeding observations	Sitting observations	Moving observations	Mean distance travelled (m)	Eagle attacks	Biting flies
0800-0900	57	308	170	200	107	3	12
0900-1000	41	367	123	209	133	1	5
1000-1100	42	410	145	187	100	4	5
1100-1200	35	400	157	196	118	4	1
1200-1300	24	298	178	241	141	1	9
1300-1400	54	364	167	168	95	2	11
1400-1500	36	321	158	218	164	2	11
1500-1600	42	408	105	225	121	0	8
1600-1700	46	452	108	198	147	1	31
1700-1800	49	338	126	211	118	0	61
r_s		0.08	-0.08	0.42	-0.50	0.119	0.63
P		NS	NS	NS	NS	NS	<0.05

The diurnal periodicity found in the occurrence of primate polyspecific associations is in agreement with studies on other African primates¹¹, and previous work on Kibale primates⁸. During the early mornings I noted that several species had frequently slept adjacent to one another. The dawn and dusk peaks in polyspecific associations probably represent the initiation and termination of night-long associations. As the *C. albigena* group slept in 45 different places on 45 different nights, there does not seem to be any shortage of sleeping sites.

These data do not support the possible existence of increased predator surveillance, increased food detection abilities or chance as causal to primate polyspecific associations. Predator detection may well be improved during the associations, but it does not

explain the temporal patterns. Feeding by two or more species in the same tree or in the same area was common, and one species may well use another as an aid in food location. But again, *C. albigena* feeding activities bore no significant relation to the temporal pattern of association. Predation, food and chance seem to do little more than add noise to a pattern that is probably due to another factor.

The importance of the significant correlation between polyspecific associations and biting fly activity is dependent on the importance of dipteran-borne disease to monkeys, and on the probability of polyspecific associations lowering the number of fly bites individuals receive. The size of the primate groups involved, and the published data on effects of number of bait animals suggest that polyspecific associations could easily reduce the number of bites received by individual monkeys. *C. albigena*, *C. ascanius*, and *P. anubis* all suffer from *Hepatozoon kochi* ('monkey malaria') which is present in Kibale¹². *C. albigena*, and *P. anubis* suffer from *Dirofilaria* sp. and *C. ascanius* have been found with microfilaria¹². All Kibale monkeys suffer from yellow fever, and other arboviruses are probably present¹². Experimental infection of three *C. ascanius* with yellow fever resulted in fevers for 1-3 d, and the death of one due to an 'intercurrent infection'¹². Such effects are likely to have considerable impact on free-living monkeys. A mechanism exists for reducing the number of fly bites received per individual monkey. The temporal pattern of this mechanism correlates with the times of day when it could function to reduce the number of fly bites per individual monkey. Disease organisms using flies as vectors are prevalent among the Kibale monkeys, and seem capable of being a strong selective factor acting to promote primate polyspecific associations.

I thank the Department of Forestry and the National Research Council, Government of Uganda for allowing this study to take place and also my field assistants for their help. The New York Zoological Society, the NIH and the NSF provided financial support.

W. J. FREELAND*

Division of Biological Sciences,
University of Michigan, Ann Arbor,
Michigan 48109

Received 24 June; accepted 15 August 1977.

*Present Address: Department of Biological Sciences, Wright State University, Dayton, Ohio 45431.

1. Freeland, W. J. *Biotropica* **8**, 12-24 (1976).
2. Willett, K. C. *Ann. trop. Med. Parasitol.* **50**, 75-80 (1956).
3. Reid, J. A. *Bull. ent. Res.* **52**, 43-62 (1961).
4. Hewitt, R. *Bird malaria* (John Hopkins, Baltimore, 1940).
5. Haddow, J. A. *Bull. ent. Res.* **33**, 91-142 (1942).
6. Struhsaker, T. T. (in the press).
7. Hamilton, W. D. J. *theor. Biol.* **31**, 295-311 (1971).
8. Struhsaker, T. T. *The Red Colobus Monkey* (University of Chicago Press, Chicago, 1975).
9. Gartlan, J. S. & Struhsaker, T. T. *J. zool. Soc. Lond.* **168**, 221-266 (1972).
10. Haddow, A. J., Gillett, J. D. & Highton, R. B. *Bull. ent. Res.* **37**, 301-320 (1947).
11. Gautier, J. P. & Gautier-Hion, A. *La Terre et la Vie* **116**, 164-201 (19xx).
12. Haddow, A. J. *Proc. zool. Soc. Lond.* **122**, 297-394 (1952).

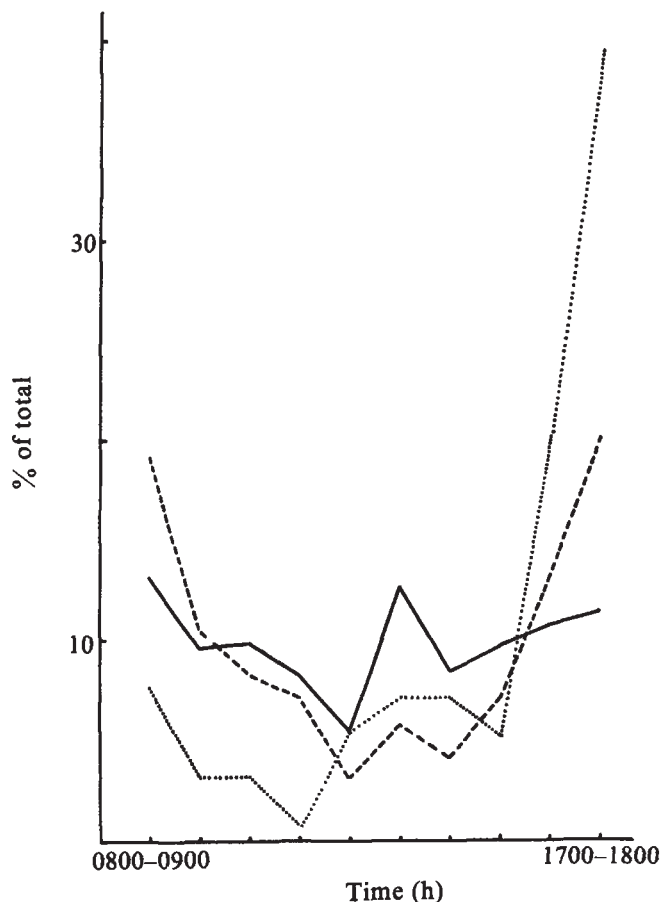


Fig. 1 *Cercopithecus albigena* polyspecific associations and the activity of Kibale biting flies and Bwamba mosquitoes. — polyspecific associations; --- Bwamba mosquitoes; Kibale biting flies.

Physiological correlates of an enzyme polymorphism

ATTEMPTS to demonstrate the action of natural selection at enzyme loci have relied mainly on *in vitro* biochemical techniques to detect differences between allelic forms of the enzymes¹⁻³. It is usually assumed that differences detected in this way will be relevant *in vivo*. The data on human inborn errors of metabolism do not support this view, however; in the majority of cases these are recessive. Heterozygotes, which usually have about 50% normal enzyme activity, rarely show clinical symptoms⁴. As any difference in activity between allozymes is likely to be much smaller, it cannot be assumed that such differences will necessarily be physiologically important. The properties of an enzyme molecule can only be biologically significant to the extent to which they affect the concentrations and flow of intermediates in its metabolic pathways. It was recently shown that changes in the concentration of most enzymes in a pathway at steady state will have little effect on flux^{5,6}. Therefore, to demonstrate that differences in fitness between strains possessing different genotypes at an enzyme locus are caused by that gene substitution, it is necessary not only to show that differences in kinetic parameters can be detected *in vitro*, but also that changes in the activity of the enzyme will affect flux through that pathway. I previously described⁷ significant changes in genotype frequency at a phosphoglucumutase locus in a wild population of field mice (*Apodemus sylvaticus*) which occurred at a time of food shortage and declining population numbers. I report here the results of some *in vivo* experiments on glycogen metabolism in animals of different genotypes at this locus, inducing glycogenolysis by fasting. The results parallel survival data from the wild population and suggest that the genotype of an individual at this locus is of physiological significance and may therefore be subject to natural selection.

The phosphoglucumutase isozyme pattern in field mice is similar to that in man⁷. Of the two loci that are active in erythrocytes, only the faster migrating one (PGM₂) is polymorphic. The two common alleles gave slow (*a*) and fast (*c*) bands on electrophoresis. This enzyme catalyses the inter-conversion of glucose-1-phosphate and glucose-6-phosphate—a reaction which links the pathways of glycogenolysis and glycolysis. In contracting muscle the flux passes directly from one pathway to the other but when liver glycogen is mobilised it is released primarily as free glucose. This pathway in liver is unusual in that it is short and has a clear beginning and end so that flux can be determined by measuring changes in stored glycogen concentration.

Liver glycogen content, estimated with the anthrone reagent^{8,9}, falls when field mice are fasted overnight. The mean concentration in 28 control animals was 28.9 mg per g wet weight (s.e., 2.41) while in 58 fasted animals it was 13.4 mg per g

(s.e., 2.41; $t_{85} = 5.4$, $P < 0.001$). PGM₂ genotype was known for 16 of the control animals. There was no significant difference between genotypes: eight animals of the *aa* genotype had a mean value of 26.1 mg per g (s.e., 3.8) while an equal number of *ac* animals had a mean of 22.2 mg per g (s.e., 3.7).

Table 1 shows the relationship between liver glycogen level and PGM₂ genotype in 48 fasted animals. The mean for the *aa* genotype is significantly lower than the combined mean for the other two groups ($t_{46} = 3.32$, $P < 0.01$). These data show a correlation between liver glycogen metabolism and PGM₂ genotype which is revealed by fasting, although genetic studies would be required to establish the causal nature of this relationship.

Table 2 shows recapture rates for each PGM₂ genotype in a natural population of *A. sylvaticus* at a time of high density, restricted food supply and falling population numbers⁷. There are significant differences in the proportion recaptured for the different genotypes and, as in Table 1, the heterozygotes give similar results to the *cc* homozygotes. Such data give a minimum estimate of survival rates¹⁰. Although errors in these estimates could arise from genetically non-uniform dispersal, they are unlikely to be large in this case as dispersal is uncommon in declining populations¹¹. The differences between genotypes in physiological responses to fasting may therefore lead to differences in survival rates at a time of food shortage.

Table 2 Recapture rate of field mice according to PGM₂ genotype

PGM ₂ genotype	<i>aa</i>	<i>ac</i>	<i>cc</i>
Total caught	22	36	18
Total recaptured	11	7	2
Proportion	0.50	0.19	0.11

Data from a natural population at Charnwood (Leicestershire, UK) giving the minimum survival rate of animals first caught in October 1974. The population was studied for 2 yr by mark-recapture trapping and gel electrophoresis⁷ and this sample represents a single cohort of non-reproductive animals. $\chi^2 = 7.32$, $P < 0.01$.

The impact of genotype at the PGM₂ locus on liver glycogen metabolism is unexpected as this locus provides less than 10% of total PGM activity in human liver¹¹. These results indicate that the two major PGM loci may have different roles *in vivo*; they have already been shown to differ in kinetic properties¹². This view is supported by several studies on the fate of labelled glucose and glucose phosphates *in vivo*¹³⁻¹⁵, which suggest that there may be two pools of glucose-6-phosphate in both liver and muscle, one of which is accessible to exogenous glucose phosphates and is involved in glycolysis while the other, from which glycogen is synthesised, is accessible only to glucose. Each of the two major PGM enzymes may be associated with only one of these glucose-6-phosphate pools, with the PGM₂ enzyme being the more important in glycogen mobilisation.

These results make it possible to visualise a physiological mechanism whereby *aa* animals are more fit than are *ac* and *cc* ones when food supply is restricted. If a higher rate of mobilisation of glycogen reserves is reflected in a higher potential for activity, *aa* individuals will have a competitive advantage when food reserves are few and widely scattered.

Considerable effort has been expended in attempts to demonstrate a selective basis for enzyme polymorphisms by means of *in vitro* biochemical studies¹⁶ although *in vitro* differences may have little physiological importance. The results of the *in vivo* experiments described here suggest that this approach may prove more rewarding.

I thank Dr J. S. Jones for discussions, Drs G. Bulfield and C. H. Langley for criticism of the manuscript and Dr R. Semeonoff for supervision. The project was supported by an SRC Research Studentship.

A. J. LEIGH BROWN*

Department of Genetics,
University of Leicester,
Leicester, UK

Table 1 Glycogen content of liver in fasted *A. sylvaticus* according to PGM₂ genotype

Genotype	<i>aa</i>	<i>ac</i>	<i>cc</i>
Mean	6.3	16.5	15.2
s.e.	1.66	4.02	2.25
N	14	18	16

Values are in mg per g wet weight. Electrophoretic methods were as described previously⁷. The genetic background of the experimental animals was standardised by taking F₁ offspring from crosses between two lines, both of which were segregating for the *a* and *c* alleles at the PGM₂ locus. Thus, all the experimental homozygotes had received one allele from each line. Food, normally freely available, was removed from the cages of fasted animals at the start of their night period. All animals were killed 15–16 h later. Liver glycogen was estimated with the anthrone reagent⁸ using sodium sulphate as a coprecipitant⁹. The recovery rate of a glycogen standard with this technique was 92% at an initial concentration of 5%. $P < 0.01$; $t_{46} = 3.32$.

Received 15 July; accepted 30 August 1977.

* Present address: Laboratory of Environmental Mutagenesis, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, North Carolina 27709.

1. Koehn, R. K. *Science* **163**, 943-944 (1969).
2. Day, T. H., Hillier, P. C. & Clarke, B. *Biochem. Genet.* **11**, 141-153; 155-165 (1974).
3. Miller, S., Percy, R. W. & Berger, E. *Biochem. Genet.* **13**, 175-188 (1975).
4. Harris, H. *The Principles of Human Biochemical Genetics* 2nd edn (North-Holland, Oxford, 1975).
5. Kacser, H. & Burns, J. A. *Symp. Soc. exp. Biol.* **27**, 65-104 (1973).
6. Rapoport, T. A., Heinrich, R. & Rapoport, S. M. *Biochem. J.* **154**, 449-469 (1976).
7. Leigh Brown, A. J. *J. Zool., Lond.* **182**, 281-289 (1977).
8. Hassid, W. Z. & Abraham, S. *Meth. Enzym.* **3**, 34-50 (1957).
9. Van Handel, E. *Analyt. Biochem.* **11**, 256-265 (1965).
10. Flowerdew, J. R. *J. Anim. Ecol.* **43**, 499-511 (1974).
11. Krebs, C. J., Gaines, M. S., Keller, B. L., Myers, J. H. & Tamarin, R. H. *Science* **179**, 35-40 (1973).
12. McAlpine, P. J., Hopkinson, D. A. & Harris, H. *Ann. hum. Genet.* **34**, 169-175 (1970).
13. Quick, C. B., Fisher, R. A. & Harris, H. *Eur. J. Biochem.* **42**, 511-517 (1974).
14. Landau, B. R. & Sims, E. A. H. *J. biol. Chem.* **242**, 163-172 (1967).
15. Dully, C. C., Bocek, R. M. & Beatty, C. H. *Endocrinology* **84**, 855-860 (1969).
16. Das, I., Sie, H.-G. & Fishman, W. H. *Archs Biochem. Biophys.* **144**, 715-722 (1971).
17. Clarke, B. *Genetics* **79**, 101-113 (1975).

Activated macrophages induce vascular proliferation

MICROVASCULAR proliferation is essential to many biological processes, such as wound healing, but the mechanisms underlying this vascular response are poorly understood¹. Neovascularisation can be induced by extracts from various cell types, including malignant solid tumour cells², normal and viral-transformed (SV40) BALB/c 3T3 cells, and diploid human embryonic lung fibroblasts³, and neutrophils⁴. Extracts of mouse salivary gland⁵ and skin⁶ have been reported to induce vascular growth, but the relevance of these observations to the vascular proliferation that occurs in wound healing and chronic inflammation is unclear. Neovascularisation is also an important component of immunological reactions. Sidky and Auerbach reported increased vessel density in the skin during graft-versus-host reactions and attributed it to the lymphocyte⁷. Herman and associates⁸ and Anderson *et al.*⁹ found an increase in capillary density, and extensive proliferation of postcapillary venular endothelium, respectively, in lymph nodes undergoing strong immunological reactions. We have shown significant endothelial proliferation in delayed hypersensitivity reactions in the skin of guinea pigs at the time of maximal mononuclear cell infiltration¹⁰ and so we have investigated whether macrophages, an important component of immunological and non-immunological inflammatory reactions, might be involved in this vascular response. We report here that macrophages activated *in vivo* and *in vitro*, and media conditioned by these cells, induce vascular proliferation in the guinea pig cornea.

Neovascularisation was assayed in the normally avascular cornea of the guinea pig eye by introducing either cells or media into the stroma and observing the growth of vascular sprouts from the limbus into the cornea. The macrophages were obtained from the peritoneal cavities of untreated guinea pigs, or from guinea pigs or mice that had previously been injected with paraffin oil or thioglycollate, respectively. We have used the terms "non-activated" and "activated" to refer to macrophages from untreated or injected animals, respectively. The activated macrophages were richer in dense phase vacuoles and in pinocytotic vesicles. Macrophages (5×10^5) were injected into the corneal stroma 1.5 mm from the limbus in 10 μ l of medium 199 (Table 1). Media conditioned by incubation with macrophages *in vitro* (legend to Table 2) were incorporated into Hydron (a polymer of hydroxyethylmethacrylate) as described by Langer and Folkman¹¹, and a 5- μ l pellet was implanted in an aseptically created pocket in the corneal stroma¹². This polymer was used because it is non-inflammatory and permits a prolonged release of incorporated compounds¹¹. Corneas were examined daily with a slit-lamp stereomicroscope ($10-40 \times$ magnification) for evidence of neovascularisation. After 7 d, just before killing, the animals were perfused intra-arterially with colloidal carbon, which filled the corneal blood vessels and provided a permanent record of the individual vascular response.

Excised corneas were fixed in buffered neutral formalin, flattened, photographed, and then processed for histological examination. A positive neovascular response was recorded when a dense brushwork of new capillary sprouts and hairpin loops grew from the adjacent limbal plexus, in a directional fashion towards the depot of test materials (Fig. 1a and b). When observed *in vivo*, proliferating capillaries were seen to grow 0.1-0.2 mm d⁻¹ and they underwent extensive branching as the response developed. Corneas with positive vascular responses did not become oedematous and histologically showed negligible infiltration by polymorphonuclear leukocytes. Responses were recorded as negative

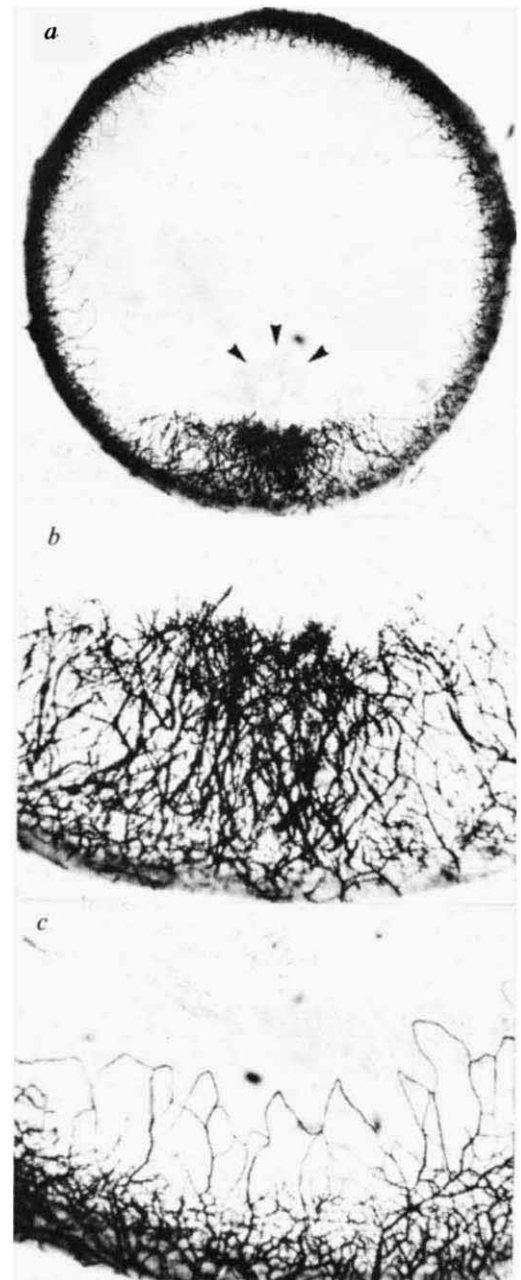


Fig. 1 Carbon-perfused vessels in Hartley albino guinea pig corneas 7 d after injection of macrophages or implantation of Hydron polymer. *a*, Positive neovascular response induced by homologous *in vivo*-activated (paraffin oil) peritoneal macrophages. Note directional ingrowth of small blood vessels from the limbal plexus adjacent to the injection site (arrows), and the normal vascular pattern around the remainder of the corneal circumference (whole mount preparation). *b*, Higher power view of positive response. The dense neovascular plexus consists of numerous small sprouts and dilated, tortuous loops, which have penetrated the corneal stroma in several planes. *c*, Typical negative response in test cornea implanted with Hydron plus unconditioned RPMI 1640 medium. Note the apparently normal limbal vessel pattern and the absence of both hairpin loops and sprouts (same magnification as *b*).

Table 1 Neovascularisation by various preparations of peritoneal macrophages injected into guinea pig corneas

Cell preparations	Total no. of corneas tested*	No. of negative	No. of positive	% Positive
Controls†	24	24	0	0
Homologous resident (non-activated)‡	12	11	1	8
Homologous activated§	18	4	14	78
Autologous activated¶	4	0	4	100
Isologous activated	8	2	6	75
Heterologous resident (non-activated)‡	12	11	1	8
Heterologous activated**	25	8	17	68
Heterologous macrophage incubated <i>in vitro</i> ††				
With latex	7	2	5	71
Without latex	4	3	1	25

*See text for criteria used in scoring neovascular responses and typical examples in Fig. 1.

†Inducing agents (paraffin oil, thioglycollate, peptone); collection media (M199); culture medium (RPMI 1640 $\pm$ 5% foetal calf serum).

‡Resident peritoneal macrophages from Hartley albino guinea pigs and BALB/c mice were obtained from normal non-stimulated animals following lavage with serum-free medium 199 (M199).

§Homologous activated macrophages were obtained from peritoneal exudates in Hartley albino guinea pigs 3 d after an intraperitoneal injection of 30 ml of sterile paraffin oil. Whole exudates suspended in M199 were layered over Lymphocyte-Separation Media (Litton Bionetics) to remove contaminating red blood cells and granulocytes, washed twice and resuspended in M199 for intracorneal injection. These preparations were 85–95% viable (Trypan blue exclusion) and consisted of 85–90% macrophages, <15% lymphocytes, <5% granulocytes, <1% other cells.

¶In four Hartley albino guinea pigs, paraffin oil-induced peritoneal macrophages obtained from individual guinea pigs were tested in their own corneas.

||Paraffin oil-activated macrophages obtained from inbred (Strain II) guinea pigs were injected in the cornea of guinea pigs of the same strain.

**Activated macrophages obtained from BALB/c mice 3 d after intraperitoneal injection of 1 ml of 3% thioglycollate were collected, processed in the same manner as guinea pig exudates and injected in Hartley albino guinea pig corneas.

†† 4×10^6 cells in 1 ml RPMI 1640 $\pm$ 5% foetal calf serum were plated in 35 $\times$ 10 mm Falcon plastic dishes and incubated at 37 °C with and without sterile latex particles (Dow Chemical, 1.09 μ m diameter, 50 beads per cell)¹⁶. After incubation for 1 h, non-adherent cells and latex were removed by washing, and adherent macrophages were detached by a 10-min treatment with 10 mM lidocaine, washed, suspended in RPMI 1640 and injected into test corneas.

when no growth, or only occasional capillary sprouts or hairpin loops, with no evidence of sustained directional growth were observed (Fig. 1c).

Tables 1 and 2 summarise the cumulative results of cells and media tested. Control materials (paraffin oil, thioglycollate, peptone; collection media, culture media incubated without cells; pellets of Hydron or Hydron and control media) showed activity in 3 out of 51 corneas tested (6%). Resident non-activated macrophages showed activity in 2 out of 24 corneas (8%). In contrast, *in vivo* activated macrophages (regardless of their source) induced a neovascular response in 41 out of 55 corneas tested (75%). In addition, macrophages stimulated *in vitro* by latex ingestion were more active (5 of 7 corneas) than unchallenged macrophages (1 of 4 corneas). Finally, incorporation of conditioned media from cultures of heterologous activated macrophages in Hydron pellets induced neovascularisation in 15 of 23 corneas (65%).

Our results demonstrate that macrophages can induce microvascular proliferation *in vivo* and they also point to some of the possible mechanisms for the reaction. First, from histological studies it seemed that the microvascular responses were not associated with an acute inflammatory response¹³. This was in contrast to other studies in which neovascularisation by macrophages was accompanied by acute inflammation¹⁴. Such a response complicates the issue of deciding which cell is

ultimately responsible for the reaction. Furthermore, in preliminary studies, we have been unable to produce corneal neovascularisation by implanting neutrophils or activated lymphocytes. Second, the neovascularisation could not be attributed solely to an immunological reaction since it was produced in a completely syngeneic cell system. These two points lead us to believe that the response is directly attributable to the macrophage, although some subtle interaction with other cell types cannot be strictly excluded. It is of interest that the responses to macrophages required their activation; namely, the macrophages either had to be obtained from peritoneal cavities following inflammation, or the cells had to be stimulated to phagocytose latex particles *in vitro*. There is ample evidence for increased macrophage function following various modes of stimulation¹⁵. Indeed, changes in phagocytic properties (for example, in secretion, in lysosomal contents) take place following stimulation of macrophages by various means. In the context of the possible *in vivo* significance of the phenomenon, it seems that the neovascularising properties of macrophages are inducible rather than constitutive. Finally, the results showing that macrophage-conditioned media contained an active material strongly suggests that the mechanisms responsible for the phenomenon may be by way of secretion of an active molecule. Macrophages can secrete a variety of macromolecules^{16–19} and characterisation of the active material is now in progress. Our

Table 2 Corneal neovascularisation induced by macrophage-conditioned media

Test material	Total no. of corneas tested	No. of negative	No. of positive	% Positive
Hydron and macrophage-conditioned media*	23	8	15	65
Hydron and control media†	12	11	1	8
Hydron carrier	15	13	2	13

*Macrophage-conditioned media were prepared by culturing peptone-induced BALB/c mouse peritoneal macrophages for 24 h in serum-free RPMI 1640 media¹⁶. After centrifugation, samples were dialysed for 72 h against RPMI 1640 to remove low molecular weight inhibitors of DNA and protein synthesis and concentrated on an Amicon UM05 filter. Conditioned medium (2.5 μ l; 0.5×10^6 – 2.5×10^6 cell equivalents) were then combined with an equal volume of sterile Hydron casting solution (prepared by dissolving the polymer in absolute alcohol at 37 °C)¹¹, and allowed to dry overnight before implantation.

†RPMI 1640 medium, incubated without cells and processed as described above.

observations therefore suggest that, among its diverse functions in the chronic inflammatory response, the macrophage may mediate microvascular proliferation.

We thank Dr S. Schlossman for his gift of Strain II guinea pigs, and Drs R. Langer and J. Folkman for assistance with the Hydron technique. This work was supported by a grant from the USPHS. M.A.G. is the recipient of an established investigatorship from the American Heart Association.

PETER J. POLVERINI

RAMZIS. COTRAN*

MICHAEL A. GIMBRONE, JR

EMIL R. UNANUE

Department of Pathology,
Peter Bent Brigham Hospital and Harvard Medical School,
721 Huntington Avenue,
Boston, Massachusetts 02115

Received 27 June; accepted 8 August 1977.

*To whom reprint requests should be addressed.

1. Folkman, J. & Cotran, R. S. *Int. Rev. exp. Path.* **16**, 207-248 (1976).
2. Folkman, J. *Cancer Res.* **34**, 2109-2113 (1974).
3. Klagsbrun, M., Knighton, D. & Folkman, J. *Cancer Res.* **36**, 110-114 (1976).
4. Fromer, C. H. & Klintworth, G. K. *Am. J. Path.* **82**, 157-167 (1976).
5. Hoffman, H., McAushan, B., Robertson, D. & Burnett, E. *Expl. Cell Res.* **102**, 269-275 (1976).
6. Wolf, J. E. Jr & Harrison, R. G. *J. invest. Derm.* **61**, 130-141 (1973).
7. Sidky, Y. A. & Auerbach, R. J. *J. exp. Med.* **141**, 1084-1100 (1975).
8. Herman, P. G., Yamamoto, I. & Mellins, H. Z. *J. exp. Med.* **136**, 697-714 (1972).
9. Anderson, N. D., Anderson, A. O. & Wyllie, R. G. *Am. J. Path.* **81**, 131-153 (1975).
10. Poverini, P. J., Cotran, R. S. & Shelley, M. M. *J. Immun.* **118**, 529-532 (1977).
11. Langer, R. & Folkman, J. *Nature* **263**, 797-800 (1976).
12. Gimbrone, M. A., Jr, Cotran, R. S., Leapman, S. B. & Folkman, J. *J. natn. Cancer Inst.* **53**, 413-427 (1974).
13. Clark, R. A. *et al. Surg. Forum* **27**, 16-18 (1976).
14. Fromer, C. H. & Klintworth, G. K. *Am. J. Path.* **81**, 531-544 (1975).
15. Unanue, E. R. *Am. J. Path.* **83**, 396-417 (1976).
16. Calderon, J., Williams, R. T. & Unanue, E. R. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4273-4277 (1974).
17. Calderon, J. & Unanue, E. R. *Nature* **253**, 359-361 (1975).
18. Calderon, J., Kiely, J. M., Lefko, J. & Unanue, E. R. *J. exp. Med.* **142**, 151-164 (1975).
19. Leibovich, S. J. & Ross, R. *Am. J. Path.* **84**, 501-513 (1976).

Human virus-infected target cells lacking HLA antigens resist specific T-lymphocyte cytotoxicity

PRODUCTS of the major histocompatibility complex (MHC) are important as target structures in T lymphocyte-mediated cytotoxicity. In mice, virus (or hapten)-sensitized T cells are cytotoxic for specific virus-infected (or hapten-modified) target cells exclusively when syngeneic or at least identical for either the K or D region of the MHC¹⁻⁵. These findings led to the 'altered-self' hypothesis⁶ which assumed that cytotoxic T cells do not recognise specific viral antigens, but rather a modification of histocompatibility antigens induced by the viral infection. One prediction of this theory was that murine cells expressing viral antigens but lacking H-2 determinants would be resistant to T cell-mediated lysis. Indeed, this fact was established recently by Zinkernagel and Oldstone using murine F 9 teratoma cells⁷. There is still no evidence for the involvement of HLA antigens in the process of immune killing of virus-infected target cells by human T lymphocytes. But, Svedmyr and Jondal have shown that peripheral T cells from patients with infectious mononucleosis (IM)—and not from normal donors—killed specifically Epstein-Barr virus (EBV) genome-carrying B lymphoblastoid cell lines⁸ and Burkitt's lymphoma cells⁹. By contrast, these T cells were devoid of any cytotoxic activity against EBV negative cell lines. Similar data were obtained by other workers^{10,11}. Thus, peripheral T cells from patients with infectious mononucleosis seemed to be sensitised to viral coded determinants, and could presumably be used in a model to study T-lymphocyte killing of virus-infected cells in human. If, in man as well as mouse, the MHC is directly involved in the target structure recognised by virus-sensitised cytolytic T lymphocytes, HLA-lacking target cells

should be resistant to immune T-cell killing. We report here that Daudi cells, originating from a human Burkitt's lymphoma line and lacking HLA products¹²⁻¹⁴, are resistant to lysis by EBV-sensitized peripheral T lymphocytes from patients with infectious mononucleosis.

The experiments involved a ⁵¹Cr release assay according to Cerottini and Brunner¹⁵ with minor modifications. Targets were cells from several B-lymphoblastoid lines and effectors were peripheral T lymphocytes from patients with infectious mononucleosis. The EBV genome-carrying lines used were T-51 (ref. 16), 8866 (ref. 17) and Daudi, and the EBV-negative line was Ramos¹⁸. All lines were maintained in our laboratory. Effector cells were purified from the blood of six patients with infectious mononucleosis, as assessed by a positive Paul Bunnell-Davidson test. Peripheral lymphocytes were isolated on Ficoll-Hypaque gradients. As unpurified peripheral blood lymphocytes exerted nonspecific killing on cell lines⁸, further purification of T cells was required. This was achieved by removal of B cells either by EAC rosetting or by use of a cytotoxic anti-B antiserum¹⁹. These processes yielded 99% pure T lymphocytes as assessed by E rosette formation.

Peripheral T lymphocytes from patients with infectious mononucleosis were cytotoxic for cells from the two HLA-positive EBV genome-carrying lines T-51 and 8866, whereas they were unable to kill cells from the HLA-positive EBV genome-negative line Ramos (Fig. 1). These results are in complete agreement with the data of Svedmyr and Jondal⁸. Strikingly, Daudi cells, though containing EBV genome, were resistant to lysis by these EBV-sensitized T cells. The human cell line Daudi lacks HLA products as assessed by the absence

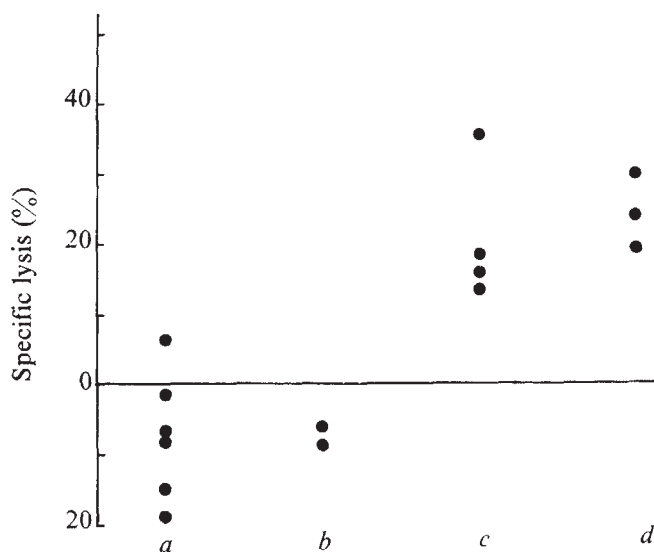


Fig. 1 Cytotoxicity of T lymphocytes from a patient with infectious mononucleosis towards target cell lines. Cytotoxicity tests were performed in U-shaped microtitre plates in a total volume of 0.2 ml RPMI 1640 medium supplemented with antibodies and 10% human AB serum (previously tested for absence of any cytotoxic activity). 2.5×10^4 labelled target cells and 1.25×10^6 or 2.5×10^6 effector cells were mixed (effector-target ratio of 50:1 or 100:1). Cultures were set up in triplicate. After 12-14 h of incubation at 37 °C in 5% CO₂ atmosphere, the supernatant of each culture was collected and ⁵¹Cr release was measured using a γ counter (Intertechnique). Simultaneously, spontaneous release in the presence of saponin was measured. Specific lysis was expressed according to the formula

$$\frac{\text{Experimental release} - \text{Spontaneous release}}{\text{Total release} - \text{Spontaneous release}} \times 100$$

a, Daudi cell line; no HLA products, EBV genome positive. b, Ramos cells, HLA phenotype 3-5-X-Y, negative for EBV genome. c, T-51 cells, HLA phenotype 1-2-8-W27, EBV genome positive. d, 8866 cells; HLA phenotype 2-3-7; EBV genome positive.

of reactivity with anti-HLA antisera tested by cytotoxicity, absorption, immunofluorescence and immunoprecipitation¹²⁻¹⁴. Daudi cells express EBV-coded determinants, however, as shown by reactivity with relevant antibodies (anti-EBNA and MA)^{20,21}. Finally, Daudi cells are susceptible to antibody-mediated immune killing, since they could be lysed in ADCC assay using anti-EBV sera and normal lymphocytes²¹ and in complement-dependant cytotoxicity assay using different anti-B sera¹³. This is also established by our own data as shown in Table 1. Daudi cells were lysed by rabbit anti-human lymphocyte serum in presence of complement and not by anti-HLA serum. Also, ADCC was positive using anti-lymphocyte and anti-B cell sera, but negative using anti-HLA serum. According to results from murine models, it is tempting to assume that the resistance to lysis by EBV-sensitised T lymphocytes exhibited by Daudi cells is due to their lack of HLA products.

Table 1 Susceptibility of Daudi cells to lysis by antibodies

Target cell treated with*	% of specific cytotoxicity Antibody- and complement- dependant* cytotoxicity	Antibody- dependant† cellular cytotoxicity (ADCC)
Normal AB serum	0.5	—
Normal rabbit serum††	0	2.9±0.6
Anti-HLA serum (Eluate from platelets)	0.5	8.5±1.2
Rabbit anti-human lymphocyte serum‡	0	20.6±1.5
Human anti-B cell serum (Eluate from Daudi cells)	ND	26.0±0.9
Normal AB serum + guinea pig complement	1.8	—
Normal rabbit serum + guinea pig complement	2	—
Anti-HLA serum + guinea pig complement	1.3	—
Rabbit anti-lymphocyte serum + guinea pig complement	96.4±1.8	—

* 5×10^4 ^{51}Cr -labelled Daudi cells were incubated with 150 μl of sera (1/30 dilution) for 30 min at 37 °C, then 100 μl of guinea pig serum (1/2 dilution) used as complement for 1 h. Cells were spun down from aliquots of cell supernatants and c.p.m. counted. % of specific cytotoxicity was calculated according to the formula

$$\frac{\text{c.p.m. total release} - \text{c.p.m. spontaneous release}}{\text{c.p.m. experimental release} - \text{c.p.m. spontaneous release}}$$

† 5×10^4 ^{51}Cr -labelled Daudi cells were incubated with the sera for 30 min at 37 °C before addition of 10^6 human lymphocytes separated by gradient of Ficoll-Trisil. Final dilution of the sera was 1/500. After 16 h, percentage of cytotoxicity was calculated as above and results expressed as compared with the effect of AB sera (% of ^{51}Cr -release in presence of experimental serum — % of ^{51}Cr -release in presence of AB serum).

‡ Absorbed on mouse cells.

This would imply that the HLA antigens are directly involved in the structure recognised by cytotoxic T lymphocytes at the surface of virus-infected target cells.

Analysis of the results of Svedmyr and Jondal⁸ shows a wide variability of the range of specific killing of each effector lymphocyte population towards different cell lines. This observation might be explained by a necessary requirement of HLA identity between target cells and effector cells, as suggested by murine models.

T lymphocytes from patients 3 and 4 did indeed exert higher specific killing for 8866 than for T-51 cells, while T lymphocytes from patient 6 were more cytotoxic for T-51 target (Table 2). This excludes a higher susceptibility to lysis of one cell line as compared with the other and may evoke a requirement of some HLA compatibility between effectors and targets. It should be noted, however, that T cells from patients 5 (HLA phenotype A.9 A.10 BW.35 BW.38) and 6 (A.9 A.19.2 B.14) were cytotoxic towards target cells with which they shared no HLA antigens.

Table 2 Cytotoxicity of T cells from individual patient with infectious mononucleosis cell lines

Patient	Target cells	Effector- target ratio	% ⁵¹ Cr release*	% specific lysis†
I	Daudi	—	32.1±2.15	—
		10:1	28.6±1.8	-5.1±2.7
		25:1	33.3±4.1	1.8±6.1
II	Daudi	—	38.5±0.1	—
		25:1	20.5±0.1	-29.2±0.15
		50:1	28.3±2.9	-16.5±4.7
III	Daudi	—	37.6±4.9	—
		50:1	25.8±1.3	-19.1±2.1
	T-51	—	27.6±1.2	—
		50:1	37.8±0.6	14±0.9
	8866	—	34.6±3.7	—
		50:1	50.5±2.2	24.3±3.4
	Ramos	—	28.7±1.3	—
		50:1	23.8±2.2	-6.8±3.1
IV	Daudi	—	37.6±4.9	—
		50:1	21.5±1.7	-25.8±2.7
		100:1	32.8±5.5	-7.8±8.8
	T-51	—	27.6±1.2	—
		50:1	27.9±1	0.9±1.9
	8866	100:1	41±5.6	22.9±10.2
	8866	—	34.6±3.7	—
		50:1	43.3±0.9	-13.4±1.4
		100:1	54.1±0.2	29.7±0.3
	Ramos	—	28.7±1.3	—
		50:1	22.5±1.1	-8.6±1.5
		100:1	22.3±1.4	-9±1.9
V	Daudi	—	26.6±0.7	—
		50:1	31.9±2.3	7.1±3.1
	T-51	—	25.5±1.5	—
VI	Daudi	50:1	36.9±1.4	15.4±2
	T-51	—	37.8±0.4	—
		50:1	39±2.2	2±3.5
	8866	100:1	37.1±6.1	-1±9.8
		—	42.6±3.3	—
		50:1	63.7±2.1	36.7±3.6
	8866	100:1	49.9±1.8	12.8±3.1
		—	37.5±1.7	—
		50:1	44.1±0.5	10.6±0.8
		100:1	48.7±4.8	17.8±7.7

* %⁵¹Cr release (±s.e.m.) calculated according to the formula

$$\frac{\text{c.p.m. experimental release} - \text{c.p.m. spontaneous release}}{\text{c.p.m. total release}} \times 100$$

† Values given ±s.e.m. methodology and specific lysis determination as in Fig. 1 legend.

Therefore in man, the histocompatibility identity did not seem to be as strongly required as in murine models.

But it is well established that series of irrelevant specificities can be detected on cell lines by usual HLA serotyping. For instance, T-51 cells additionally reacted with sera directed against HLA 10 and 21 and 8866 with anti-HLA 1-11-28-19-2-27-5-29-35-21 antisera. These irrelevant reactions could be due to alloantibodies directed to Ia-like B-cell antigens. It might be postulated that the expression of HLA antigens at the surface of cell lines allows a high degree of cross reactivity with alloantisera²², or that virus-infected cell possibly express new HLA specificities, as suggested in the mouse²³. If the latter were so, this would explain why cell lines would not apparently be submitted to strong allogeneic restriction when used as target cells. It remains possible that much higher specific killing could be observed if the target cells were autologous EBV-infected B lymphocytes from the infectious mononucleosis patient whose T cells are used as effectors. This possibility is presently under investigation.

But, regardless of the apparent lack of allogeneic restriction in this model, the fact that Daudi cells are resistant to lysis by EBV-sensitised T cells is the first evidence to suggest direct HLA involvement in anti-viral T-cell killing in man.

We thank Mrs Colette Chereau, Maryvonne Guillard, Marielle Duthail and Yvonne Thies for technical assistance. This work was supported by grants from Inserm and Institut Gustave-Roussy, Villejuif, France.

Note added in proof: After submission of our paper, another example of implication of HLA in the target structure of T-cell killing was published (Goulmy, E., Termijtellen, A., Bradley, B. A. and van Rood, J. J. *Nature* **266**, 544-546, 1977) concerning the response to the H-4 antigens. In this case the authors found an allogeneic restriction to HLA-2.

T. TURSZ

Laboratoire d'Immuno-hematologie et Service de Medecine,
Institut Gustave Roussy, 94800 Villejuif, France

W. H. FRIDMAN

A. SENIK

Laboratoire d'Immunologie Cellulaire,
IRSC, 94800 Villejuif, France

A. TSAPIS

Laboratoire d'Immunochimie et d'Immunopathologie
(Inserm U108),
Institut de Recherches sur les Maladies du Sang,
Hôpital Saint-Louis

M. FELLOUS

Laboratoire de Genetique de la Transplantation Humaine
(Inserm U93),
Institut de Recherches sur les Maladies du Sang,
Hôpital Saint-Louis, 75010 Paris, France

Received 15 March; accepted 26 August 1977.

1. Zinkernagel, R. M. & Doherty, P. C. *Nature* **248**, 701-702 (1973).
2. Koszinowski, U. & Ertl, H. *Nature* **255**, 522-524 (1975).
3. Shearer, G. M. *Eur. J. Immun.* **4**, 527-533 (1974).
4. Schrader, J. W. & Edelman, G. M. *J. exp. Med.* **143**, 601-614 (1976).
5. Burakoff, S. J., Germain, R. N., Dorf, M. E. & Benacerraf, B. *Proc. natn. Acad. Sci. U.S.A.* **73**, 625-629 (1976).
6. Doherty, P. C., Blanden, R. V. & Zinkernagel, R. M. *Transplant. Rev.* **29**, 89-124 (1976).
7. Zinkernagel, R. M. & Oldstone, M. B. A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3666-3670 (1976).
8. Svedmyr, E. & Jondal, M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1622-1626 (1975).
9. Jondal, M., Svedmyr, E., Klein, E. & Singh, S. *Nature* **255**, 405-407 (1975).
10. Hutt, L. M., Huang, Y., Dascomb, H. E. & Pagano, J. S. *J. Immun.* **115**, 243-248 (1975).
11. Royston, I., Sullivan, J. L., Periman, P. O. & Perlin, E. *New Engl. J. Med.* **293**, 1159-1163 (1975).
12. Fellous, M., Mortchelewicz, F., Kamoun, M. & Dausset, J. in *Histocompatibility Testing* (ed. Kissmeyer-Nielsen) 708-711 (Munksgaard Copenhagen, 1975).
13. Jones, E. A., Goodfellow, P. N., Bodmer, J. G. & Bodmer, W. F. *Nature* **256**, 650-653 (1975).
14. Reisfeld, R. A., Seviess, E. D., Pellegrino, M. A., Ferrone, S. & Poulik, M. D. *Immunogenetics* **2**, 184-197 (1975).
15. Cerrotini, J. C. & Brunner, K. T. in *In Vitro Methods in Cell Mediated Immunity* (eds Bloom, B. R. & Glade, P. R.) 369, (Academic, New York and London, 1971).
16. Orkin, S. H., Buchanen, P. D., Ryount, W. G., Reisner, H. & Littlefields, G. W. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2401-2405 (1973).
17. Pious, D., Bodmer, G. J. & Bodmer, W. F. *Tissue Antigens* **4**, 247-256 (1974).
18. Steinitz, M. & Klein, G. *Virology* **70**, 570-573 (1976).
19. Tursz, T., Preudhomme, J. L., Labaume, S., Matuchansky, C. & Seligmann, M. *J. clin. Invest.* **60**, 405-410 (1977).
20. Klein, G., Wiener, F., Zech, L., Zur Hausen, H. & Reedman, B. *Int. J. Cancer* **14**, 54-64 (1974).
21. Jondal, M. *Clin. exp. Immun.* **15**, 1-5 (1976).
22. Dick, H. M., Bodmer, J. G., Steel, C. M., Crichton, W. B. & Evans, J. in *Histocompatibility Testing* (ed. Kissmeyer-Nielsen) 671-676 (Munksgaard, Copenhagen, 1975).
23. Garrido, F., Schirmacher, V. & Festenstein, H. *Nature* **259**, 228-229 (1976).

killing can be associated with either *K* or *D* region determinants, that is, cytotoxic T cells from virus-infected mice will kill virus-infected target cells which share either the *K* or *D* region with the infected strain². Our studies, however, indicate that in the FV system there is a selective participation of *H-2* determinants in virus-specific killing by cytotoxic T cells: BALB.B mice immunised with cultured, syngeneic FV-induced tumour cells produce T cells with a cytotoxicity for other FV-induced tumour cells which depends only on identity at the *H-2D* (and not the *H-2K*) region of the *H-2^b* haplotype. This *H-2D^b* restriction in FV-specific cytotoxicity correlates with previous observations concerning: (1) partial co-capping of *H-2D^b* and not *H-2K^b* specificities following treatment with anti-FV antiserum³; (2) selective incorporation of *H-2D^b* and not *H-2K^b* specificities into mature FV particles⁴; and with our present observation of (3) blocking of T-cell cytotoxicity by anti-*H-2D^b* and not anti-*H-2K^b* antisera. We suggest that the basis for this correlation is the formation of *H-2D^b*/viral protein complexes on the surfaces of FV-infected cells, forming an 'altered self' molecule recognised by cytotoxic T cells as previously proposed for the mechanism of *H-2* restriction⁵.

Mice for these studies were of the BALB/c strain (*H-2^d*) or of congenic strains differing at the *H-2* complex: BALB.B (*H-2^b* from C57BL/10) or BALB.G (*H-2^a*, a recombinant haplotype originating from an *H-2^b*/*H-2^d* heterozygote, from HTG/Go). Cultured FV-induced tumour cell lines derived from mice of these strains were of the HFL series⁶: HFL/b from BALB.B, HFL/d from BALB/c, and HFL/g from BALB.G. Cells of each of these lines express both the FV-induced FMR cell-surface antigen and the virus envelope antigens associated with the gp70 protein of FV.

BALB/c, BALB.B and BALB.G mice were immunised by intraperitoneal inoculation of syngeneic HFL tumour cells. Peritoneal exudate cells (PEC) taken on day 5 after the second or third inoculation were assayed for lymphocyte-mediated cytotoxicity (LMC) by the method of Burke *et al.*⁷, using ⁵¹Cr-labelled HFL tumour cells as targets. Controls for the sensitivity of the LMC reaction to pretreatment with anti-Thy-1.2 antiserum and complement were included in most experiments and demonstrated that the killing was attributable to T cells.

Table 1 *H-2* restriction of T-cell cytotoxicity

Mice immunised*	<i>H-2</i> haplotype	Syngeneic immunising cells	Target cell cytotoxicity (% ⁵¹ Cr release)			
			HFL/b	HFL/g	HFL/d	HFL/k
BALB.B	b b b b	HFL/b	18	15	0	0
BALB.G	d d d b	HFL/g	20	37	1	0
BALB/c	d d d d	HFL/d	0	0	0	0

In the LMC, 2.5×10^6 immune PEC in 1 ml and 5×10^4 ⁵¹Cr-labelled tumour cells in 0.1 ml (PEC-target cell ratio, 50:1) were mixed and incubated at 37 °C in 5% CO₂ atmosphere for 6 h. The cells were suspended in Medium 199 plus 20% foetal calf serum. Specific release of radioactivity into supernatant was compared with that from tumour cells alone frozen and thawed four times.

*PEC were collected at previously determined peaks of activity, 5 d after the second or third immunisation.

Evidence for an *H-2*/viral protein complex on the cell surface as the basis for the *H-2* restriction of cytotoxicity

CYTOTOXIC T lymphocytes generated in mice by immunisation with syngeneic Friend virus (FV)-induced tumour cells show *H-2* restriction, that is, their cytotoxicity is specific not only for FV-associated antigenic determinants expressed on the surfaces of potential target cells but also for products of genes associated with the major histocompatibility complex, *H-2* (ref. 1). We report here the results of studies concerning the role of molecules governed by genes in the *K* and *D* regions of the *H-2* complex in the stimulation of FV-specific immune responses. Studies in other viral systems have shown that *H-2*-restricted T-cell

One criterion used to define the extent of *H-2* identity with target cells which would permit the T cells to exert their cytotoxicity was direct cross reactivity in the LMC assay (Table 1). The marked cross reactivity of BALB.B anti-HFL/b and BALB.G anti-HFL/g PEC indicated that *H-2D^b* identity was sufficient to permit killing to occur. Whether *H-2K^b* identity was also sufficient could not be tested by this technique, since no target cell was available which shared this genotype but differed in the *H-2D* region. The observation that BALB.G anti-HFL/g PEC were ineffective against HFL/d tumour cell targets suggested that *H-2K^d*, *H-2I^d* and *H-2S^d* region identity was not sufficient to permit killing.

Another criterion used to elucidate the precise requirements for $H-2$ identity between immunising and target cells, was the capacity of high concentrations of specific anti- $H-2$ antisera added to the reaction mixture to block T-cell killing of syngeneic target cells⁸. Accordingly, we tested the capacity of antisera specific for either $H-2K^b$ or $H-2D^b$ antigenic determinants to inhibit the cytotoxicity of BALB.B anti-HFL/b PEC for HFL/b target cells. The results (Fig. 1) showed that anti- $H-2D^b$ antibodies were effective in blocking the reaction whereas anti- $H-2K^b$ antibodies showed no greater inhibitory effect than did irrelevant anti- $H-2^k$ antibodies. Preliminary attempts to detect blocking in this same reaction by antibodies to specific FV proteins met with little success; of the various antisera tested, only one of three available pools of goat anti-gp70 antiserum demonstrated significant inhibitory activity (data not shown).

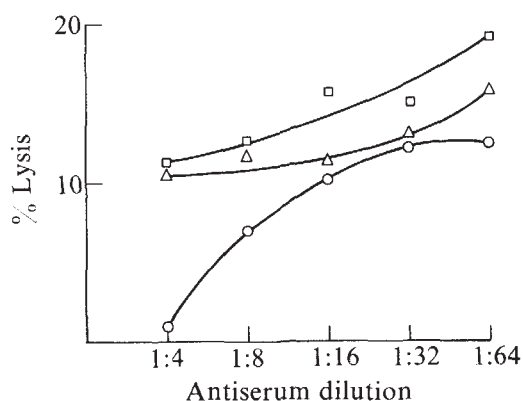


Fig. 1 Lysis of HFL/b target cells by peritoneal exudate cells (PEC) from BALB.B mice immunised with HFL/b tumour cells. PEC (2.5×10^6) in 0.2 ml and 5×10^4 ^{51}Cr -labelled tumour cells in 0.1 ml were mixed together with 0.1 ml of diluted anti- $H-2$ antiserum and assayed for lymphocyte-mediated cytotoxicity as noted in Table 1. Antisera used were: $\circ$, anti- $H-2D^b$ (BALB/c anti-BALB.G normal spleen cells); $\triangle$, anti- $H-2K^b$ [(A $\times$ BALB.G) F_1 anti-BALB.B normal spleen cells]; $\square$, anti- $H-2^k$ (BALB.B anti-BALB.K normal spleen cells).

These experiments indicate that T-cell recognition of foreign antigenic determinants on FV-induced $H-2^b$ and $H-2^d$ tumour cells involved specificities associated with both the $H-2D^b$ molecule and a virus molecule, perhaps gp70. The $H-2K^b$ and $H-2K^d$ molecules did not seem to be involved in this recognition, and no assessment of the possible involvement of the $H-2D^d$ molecule could be made, since cytotoxic activity has never been detected in any of several experiments with PEC from BALB/c mice immunised with HFL/d tumour cells. Our findings correlate with those of recent studies in which disrupted FV particles collected from the serum of infected (BALB/c $\times$ BALB.B) F_1 mice ($H-2^b/H-2^d$) were examined for $H-2$ antigen content and found to have incorporated $H-2D^b$ but not $H-2K^b$, $H-2D^d$ or $H-2K^d$ determinants⁴. These two sets of experiments strongly imply that a physical association between a viral molecule and the $H-2D^b$ molecule occurs on cell surfaces which is based on an affinity between the two molecules which does not exist between the viral molecule and the $H-2K^b$, $H-2D^d$ or $H-2K^d$ molecules.

Two further sets of experiments provide an indication that this molecular association may be of importance *in vivo*. First, the rejection of grafts of cultured, syngeneic HFL cells was correlated with the presence of $H-2D^b$ molecules (Table 2). All BALB.B mice and 76% of BALB.G mice rejected HFL/b and HFL/g cell grafts, respectively, whereas only 8% of BALB/c mice rejected HFL/d cell grafts.

Second, we attempted to demonstrate that the antigenic specificities of cultured HFL tumour cells recognised by immune T cells were the same as those of primary FV-infected cells *in vivo* by carrying out adoptive transfer experiments (Table 3). Spleen

Table 2 Survival of mice challenged with syngeneic tumour cells

Recipient mice*	Tumour cells	Cell dose	No. survivors/no. inoculated
BALB.B	HFL/b	5×10^6	57/57
BALB.G	HFL/g	5×10^6	16/21
BALB/c	HFL/d	5×10^6	1/12
		5×10^5	2/14

*Mice were inoculated i.p. with syngeneic tumour cells and observed for 8 weeks.

cells from BALB.B and BALB.G donor mice which had rejected syngeneic HFL cell grafts were injected into irradiated (350 R) non-immune syngeneic recipient mice which then received a high dose of FV. Control recipients of non-immune spleen cells developed the pronounced splenomegaly characteristic of the FV disease, whereas recipients of immune spleen cells showed no splenomegaly. Pretreatment of the immune donor spleen cells with anti-Thy-1.2 antiserum and complement abolished their capacity to transfer immunity to the disease.

Table 3 Adoptive transfer of immunity to FV by syngeneic T cells from mice immunised with syngeneic HFL tumour cells

Irradiated recipients*	Mean spleen weights $\pm$ s.d. (g)		
	A†	B	C
BALB.B	0.15 ± 0.01	2.31 ± 0.27	
BALB.G	0.10 ± 0.01	1.41 ± 0.21	2.33 ± 0.07

*Recipient mice were irradiated (350 R) and inoculated with 1.5×10^7 syngeneic spleen cells i.v.; 24 h later the mice received 3,000 FFU of FV i.v. Spleen weights were determined after 30 d or earlier if mice appeared moribund. Each group included 3–5 mice.

†Group A received immune syngeneic spleen cells pretreated with normal mouse serum and complement; group B received immune syngeneic spleen cells pretreated with anti-Thy-1.2 serum plus complement; group C received syngeneic spleen cells from unimmunised mice.

FV-infected cells which express the $H-2^b$ or $H-2^d$ but not the $H-2^k$ haplotypes show cross-reacting virus-associated antigenic specificities capable of eliciting a cytotoxic T-cell response directed against specificities associated with the $H-2D^b$ molecule. This immunogenicity seems to be the result primarily of the formation on the cell surface of an $H-2D^b$ /viral molecule complex, but the capacity to form such a complex is not, in general, a property of other $H-2D$ or of $H-2K$ molecules. It seems likely that this phenomenon is, at least in part, the basis of the observation⁹ that $H-2$ type is a major determinant of susceptibility or resistance to viral leukaemogenesis in mice.

This work was supported by a contract with the Virus-Cancer Program and by a grant, both from the National Cancer Institute. K.J.B. was supported by a training grant from the National Cancer Institute. We acknowledge the gift of some of the anti-MuLV protein antisera used from Drs M. Strand and J. T. August.

KENNETH J. BLANK

FRANK LILLY

Department of Genetics,
Albert Einstein College of Medicine,
Bronx, New York 10461

Received 21 July; accepted 22 August 1977.

- Blank, K. J., Freedman, H. A. & Lilly, F. *Nature* **260**, 250–252 (1976).
- Doherty, P. C. & Zinkernagel, R. M. *J. exp. Med.* **141**, 502–508 (1975).
- Bubbers, J. E., Steeves, R. A. & Lilly, F. *Proc. Am. Ass. Cancer Res.* **17**, 93 (1976).
- Bubbers, J. E. & Lilly, F. *Nature* **266**, 458–459 (1977).
- Zinkernagel, R. M. & Doherty, P. C. *Nature* **248**, 701–702 (1974).
- Freedman, H. A. & Lilly, F. *J. exp. Med.* **142**, 212–223 (1975).
- Berke, G., Sullivan, K. A. & Amos, D. B. *J. exp. Med.* **135**, 1334–1348 (1972).
- Shrader, J. W. & Edelman, G. M. *J. exp. Med.* **143**, 601–614 (1976).
- Lilly, F. *J. exp. Med.* **127**, 465–473 (1968).

Effect of protease inhibitors and substrates on deoxycorticosterone binding to its receptor in dog MDCK kidney cells

The dog kidney MDCK cell line contains a protein in its cytosol that specifically binds deoxycorticosterone (DOC) and can concentrate DOC in its nuclei. A Scatchard analysis of ^3H -DOC binding shows that the K_d is about $8 \times 10^{-8} \text{ M}$ (refs 1, 2). We describe here a study of the effects of various protease inhibitors on DOC binding to its receptors. The DOC receptor in MDCK cells may contain a specific structure which recognises protease inhibitors.

The protease inhibitors, phenylmethyl sulphonylfluoride³ (PMSF), tosyl-lysine chloromethyl ketone⁴ (TLCK) and tosylamide-phenylethyl-chloromethyl ketone⁵ (TPCK) eliminate specific H-DOC binding activity in the MDCK cytosol. This occurs both when these compounds are present in the homogenising medium and when added to the 104,000g cytosol just before ^3H -DOC. The elimination of ^3H -DOC binding could be due to these protease inhibitors either interacting with the DOC receptor or preventing a proteolytic process necessary for activation of pro-receptor for DOC. To distinguish between these possibilities, we incubated the 104,000g cytosol with $3 \times 10^{-8} \text{ M}$ ^3H -DOC for 2.5 h at 0 °C. By this time, equilibrium binding of the ^3H -DOC has been achieved. Then a 1,000-fold molar excess of nonradioactive DOC, 1 mM PMSF, 1 mM TLCK, or 0.5 mM TPCK was added to aliquots of the cytosol. In 1 h the three protease inhibitors eliminated specific ^3H -DOC binding; whereas the binding of ^3H -DOC by the untreated cytosol did not change (Table 1). We therefore conclude that these protease

Table 2 Dependence of inhibition of ^3H -DOC binding on the concentration of protease inhibitors

Compound	Concentration (M)	% Inhibition
TPCK	1.5×10^{-6}	97
	1.5×10^{-7}	88
	3.0×10^{-8}	38
PMSF	1.5×10^{-6}	50
	1.5×10^{-6}	34
TLCK	3×10^{-4}	47
	1.5×10^{-5}	28

MDCK cell cytosol (with 4 mM dithiothreitol in the homogenising buffer) was incubated with $3 \times 10^{-8} \text{ M}$ ^3H -DOC for 1.5 h. Aliquots were then incubated for 2.5 h with $6 \times 10^{-6} \text{ M}$ DOC, or the above compounds. Bound steroid was separated from unbound steroid on Sephadex G-25 columns. Per cent inhibition refers to the inhibition of specifically bound counts, determined by subtracting the amount of ^3H -DOC bound in presence of $6 \times 10^{-6} \text{ M}$ DOC from the amount of ^3H -DOC bound in the absence of $6 \times 10^{-6} \text{ M}$ DOC.

inhibitors interact directly with the DOC receptor. Since DOC binding was inhibited by concentrations of protease inhibitors that were several orders of magnitude greater than the concentration of ^3H -DOC (Table 1), we conducted more detailed studies at lower concentrations of inhibitor to determine if inhibition of DOC binding occurred at concentrations which would be indicative of stoichiometric reaction with the DOC receptor. We also sought to determine differences in the affinity of these compounds for the DOC receptor. As shown in Table 2, TPCK apparently has a high affinity for the DOC receptor, since only a fivefold excess concentration inhibits DOC binding by nearly 90%. Furthermore, PMSF, TPCK, and TLCK all have different affinities for the DOC receptor in MDCK cells.

Such inhibition of DOC binding by protease inhibitors could be due to interaction either with a specific structure in the DOC receptor that recognises protease inhibitors or of a nonspecific type that denatures the DOC receptor. One way in which TPCK or TLCK might denature the DOC receptor would be to bind the thiol groups⁶⁻⁸ on the receptor. This possibility is unlikely, however, since TLCK and TPCK can inhibit ^3H -DOC binding to MDCK cytosol in the presence of 4 mM dithiothreitol and since TPCK is effective at concentrations as low as 10^{-7} M .

Additional observations with protease substrates also support our postulate that the DOC receptor in MDCK cells contains a specific structure that recognises the protease inhibitors. Moreover, this specific structure may be similar to that recognised by the protease inhibitors in proteases. Table 3 shows that four protease substrates, *p*-toluenesulphonyl-L-arginine methyl ester (TAME), L-tryptophan methyl ester (TME), acetyl-L-tyrosine methyl ester (ATEE), and benzoyl-L-arginine ethyl ester (BAEE) inhibit ^3H -DOC binding to MDCK cell cytosol. A more detailed examination of the relationship of the structure of the substrate and its effectiveness as an inhibitor of DOC

Table 1 Effect of protease inhibitors on ^3H -DOC binding in MDCK cell cytosol

Sample	Bound ^3H -DOC (c.p.m. per mg protein)*
^3H -DOC ($3 \times 10^{-8} \text{ M}$) 2.5 h incubation	220,300
^3H -DOC ($3 \times 10^{-8} \text{ M}$) + DOC ($3 \times 10^{-8} \text{ M}$) 2.5 h incubation	19,300
^3H -DOC ($3 \times 10^{-8} \text{ M}$) 3.5 incubation	214,900
^3H -DOC ($3 \times 10^{-8} \text{ M}$) + DOC ($3 \times 10^{-8} \text{ M}$) 3.5 h incubation	19,900
^3H -DOC ($3 \times 10^{-8} \text{ M}$) 2.5 h incubation, then	
PMSF (1 mM) 1 h	21,600
TLCK (1 mM) 1 h	22,000
TPCK (0.5 mM) 1 h	20,000
DOC ($3 \times 10^{-5} \text{ M}$) 1 h	19,800

MDCK cells were cultured in Dulbecco-Vogt modified Eagles medium with 10% foetal calf serum on 100-mm Falcon tissue culture dishes. Two days after cell growth had stopped, the cells were washed once with cold phosphate-buffered saline and once with the homogenising solution, 10 mM Tris pH 7.4 + 2 mM EDTA. The cells were removed from the dish with a rubber policeman and homogenised in a Dounce homogeniser. The homogenate was centrifuged at 104,000g for 1 h. The resulting cytosol was incubated at 0 °C with $3 \times 10^{-8} \text{ M}$ ^3H -DOC (48 Ci mmol, New England Nuclear). An aliquot was also incubated with $3 \times 10^{-8} \text{ M}$ ^3H -DOC + $3 \times 10^{-5} \text{ M}$ non-radioactive DOC. After 2.5 h at 0 °C, aliquots from the $3 \times 10^{-8} \text{ M}$ ^3H -DOC sample were incubated with 1 mM PMSF, 1 mM TLCK, 0.5 mM TPCK, or $3 \times 10^{-5} \text{ M}$ non-radioactive DOC for 1 h at 0 °C. Bound steroid was separated from unbound steroid on BioGel P-10 columns (1 × 10 cm). Samples were done in duplicate and the results averaged. Variation was less than 10%. Fractions were counted in a liquid scintillation cocktail containing Omnifluor (New England Nuclear), Triton X-100 and toluene. Thin-layer chromatography, 100 min using dichloromethane-acetone (80:20) as the developing solvent, shows that incubation of ^3H -DOC with PMSF, TPCK or D-TME does not alter the chromatographic behaviour of ^3H -DOC.

*220,300 c.p.m. per mg = $1.26 \times 10^{-11} \text{ mol}$ per mg protein.

Table 3 Effect of protease substrates on ^3H -DOC binding

Substrate	% Inhibition
1 mM L-TAME	88
1 mM D-TAME	95
10 mM L-TME	73
10 mM D-TME	78
1 mM D-TME	25
1 mM L-ATEE	43
5 mM L-BAEE	60

Inhibition of binding of $3 \times 10^{-8} \text{ M}$ ^3H -DOC to MDCK cytosol by protease substrates was measured as described in Table 1, except that 4 mM dithiothreitol was included in the homogenising buffer.

binding was conducted for arginine and derivatives (Table 4). D-TAME produced 76% inhibition at 1.5×10^{-6} M, while 6×10^{-5} M tosyl-L-arginine was required to produce 46% inhibition and 10^{-2} M L-arginine produced only 17% inhibition. These findings emphasise the importance of the ester structure in effecting inhibition of DOC binding.

We then determined if protease substrates and inhibitors reversibly inhibit ^3H -DOC binding to its receptor. MDCK cytosol was incubated with 3×10^{-8} M ^3H -DOC and D-TME to eliminate specific ^3H -DOC binding. The inhibited cytosol was filtered through a Sephadex G-25 column to remove unbound D-TME, and the filtrate could then bind ^3H -DOC (Table 5). Similar results were found for inhibition of ^3H -DOC binding to MDCK cell cytosol by D-TAME. We conclude that the inhibition of ^3H -DOC binding to MDCK cytosol by D-TME and D-TAME is reversible. In contrast, a similar experiment with the protease inhibitor TPCK showed that the inhibition of DOC binding by TPCK is not reversible (Table 5). This suggests that TPCK binds covalently to the DOC receptor, as would be expected since TPCK is a covalent inhibitor of chymotrypsin.

Table 4 Dependence of inhibition of ^3H -DOC binding on the concentration of protease substrates

Compound	Concentration in cytosol (M)	% Inhibition
L-TAME	1.5×10^{-5}	55
D-TAME	6×10^{-6}	95
	1.5×10^{-6}	76
Tosyl-L-arginine	1.5×10^{-4}	84
	6×10^{-5}	46
L-Arginine	1.0×10^{-2}	17
	5×10^{-3}	3

MDCK cell cytosol (with 4 mM dithiothreitol in the homogenising buffer) was incubated with 3×10^{-8} M ^3H -DOC for 1.5 h. Aliquots were then incubated for 2.5 h with 6×10^{-6} M DOC, or the above compounds. Bound steroid was separated from unbound steroid on Sephadex G-25 columns. Per cent inhibition refers to the inhibition of specifically bound counts, determined by subtracting the amount of ^3H -DOC bound in presence of 6×10^{-6} M DOC from the amount of ^3H -DOC bound in the absence of 6×10^{-6} M DOC.

Although the details of the mechanism of inhibition of ^3H -DOC binding are not known, several structural features might be anticipated from the well understood mechanism of inhibition of proteases by PMSF, TLCK and TPCK^{3-6,9}. PMSF inhibits a broad class of proteases by binding covalently to the active serine residue, while TLCK and TPCK inhibit by binding covalently to a histidine residue spatially adjacent to the serine. Since both protease substrates and inhibitors prevent DOC binding, it is possible that a structure similar to that of the serine proteases is present in the DOC receptor. The DOC receptor is unique, however, in binding both a trypsin substrate (TAME) and a chymotrypsin inhibitor (TPCK). This might be expected for a receptor that recognises a steroid nucleus rather than specific amino acid residues.

An important question is whether protease inhibitors and substrate recognise only the DOC receptor or a structural feature common to several steroid binding proteins. Our evidence suggests that the inhibition of steroid binding by protease inhibitors and substrates is a more general phenomenon. For example, PMSF, TLCK, TPCK, TME, and TAME inhibit the binding of ^3H -oestradiol to purified rat α foetoprotein (M.B. and D.F., in preparation). In addition, PMSF, TLCK, and TPCK inhibit the binding of ^3H -dexamethasone and ^3H -oestradiol

to rat kidney cytosol (M.E.B., N. Anderson, D. Vaughn, and D.F., unpublished results).

From our results, we cannot indicate whether the steroids and protease inhibitors and substrates bind to the same active site in the steroid receptor or if there is a specific catalytic function associated with the ability of steroid receptors to bind protease inhibitors and substrates.

Table 5 Reversibility of inhibition of DOC binding

Condition	% Inhibition by TME (15 mM)	% Inhibition by TPCK (1.5 μM)
a, Inhibitors present with ^3H -DOC-receptor complex for 1 h*	92	100
b, Inhibitors in (a) removed by filtration on G-25 Sephadex; followed by re-equilibration with ^3H -DOC for 2 h†	9	95

*MDCK cell cytosol was incubated with 3×10^{-8} M ^3H -DOC for 2 h then aliquots were incubated for 1 h with 15 mM D-TME or 1.5×10^{-6} M TPCK. Bound steroid was separated from unbound steroid on Sephadex G-25 columns.

†Aliquots of 1 ml cytosol which had been incubated with 3×10^{-8} M ^3H -DOC for 3 h with or without 15 mM D-TME or 1.5×10^{-6} M TPCK for 1 h in (a) were filtered on Sephadex G-25 columns to remove unbound ^3H -DOC, D-TME, or TPCK. Then the filtrate was incubated with 3×10^{-8} M ^3H -DOC with or without 6×10^{-6} M DOC for 2 h. Bound steroid was separated from unbound steroid on Sephadex G-25 columns. % Inhibition refers to the inhibition of specifically bound counts, which were determined by subtracting the amount of ^3H -DOC bound in the presence of 6×10^{-6} M DOC from the amount of ^3H -DOC bound in the absence of 6×10^{-6} M DOC.

Nevertheless, two important applications of our results should be noted. First, the reversible inhibition of steroid binding by the D forms of TAME and TME suggest that it may be possible to purify steroid receptors by using the protease substrates as the ligand in affinity chromatography. To date, most attempts at purification of steroid receptors by affinity chromatography using steroid ligands have been unsuccessful, apparently because the high affinity of the steroid ligand for the receptor makes difficult the elution of the receptor in an active state or in high yield^{10,11}. With D-TAME or D-TME, however, the lower affinity of the steroid receptor for the protease substrate could improve the yield of receptor. Second, a new class of compounds which could react with steroid receptors *in vivo* is suggested. PMSF, TPCK, and TLCK enter cells, covalently, interact with proteins, and are unstable at pH 7.4 (37 °C)¹². Therefore, inhibitors could react with steroid receptors regionally with little toxicity or effect in areas of the organism distant from the point of application. Such compounds, or more selective protease inhibitors, might prove useful in treating localised steroid dependent tumours. The ability of protease inhibitors and substrates to block steroid binding to receptors may thus provide information on both structure and function of the steroid receptors.

We thank D. Vaughn for assistance and Dr N. Anderson for discussion of his model of formation of steroid receptors from proreceptors in cell cytosol. This work was supported by grants from the USNIH (USPHS AM 14915 and USPHS CA 15796).

MICHAEL E. BAKER
DARRELL D. FANESTIL

Division of Nephrology,
Department of Medicine, M-013,
University of California, San Diego,
La Jolla, California 92093

Received 7 October 1976; accepted 1 August 1977.

- 1 Ludens, J. H., Vaughn, D. A., Mawe, R. C., Saier, M. H. & Fanestil, D. D. *Fedn Proc.* 34, 663 (1975).
- 2 Ludens, J. H., Vaughn, D. A., Mawe, R. C. & Fanestil, D. D. *J. Steroid Biochem.* (in the press).
- 3 Fahrney, D. E. & Gold, A. M. *J. Am. Chem. Soc.* 85, 997-1000 (1963).
- 4 Shaw, E. M., Mares-Guia & Cohen, W. *Biochemistry* 4, 2219-2224 (1965).
- 5 Schoellmann, G. & Shaw, E. *Biochemistry* 2, 252-255 (1963).
- 6 Shaw, E. M. in *The Enzymes* 1, 3rd ed. (ed. Boyer, P. D.) 91-146, (Academic, New York, 1970).
- 7 Stein, M. J. & Liener, J. E. *Biochem. biophys. Res. Commun.* 26, 376-382 (1967).
- 8 Bender, M. L. & Brubacher, L. J. *J. Am. Chem. Soc.* 88, 5880-5889 (1966).
- 9 Shaw, E. M. in *Proteases and Biological Control* (eds Reich, E., Rifkin, D. B. & Shaw, E.) 455-465 (Cold Spring Harbor Laboratory, New York, 1975).
- 10 Ludens, J. H., DeVries, J. R. & Fanestil, D. D. *J. biol. Chem.* 247, 7533-7538 (1972).
- 11 Kuhn, R. W., Schrader, W. T., Smith, R. G. & O'Malley, B. W. *J. biol. Chem.* 250, 4220-4228 (1975).
- 12 Gold, A. M. *Meth. Enzym.* 2, 706-711 (1967).

Non-sequential expression of multiple immunoglobulin classes by isolated B-cell clones

As humoral immunity to a given antigen is generally thought to display an early IgM response which is later dominated by IgG synthesis, it has long been thought that clones of antibody-producing cells switched from μ to γ (refs 1-5). The sequence of appearance of immunoglobulin classes produced by the progeny of isolated precursors (B cells) to antibody-forming cells may answer the questions (1) does a clone of cells switch from μ to γ to α , and (2) do all clones switch in the same order following antigen stimulation? Results presented here show that there is no order of expression of immunoglobulin class with time by B-cell clones stimulated by antigen in the presence of carrier help, because any class can appear first or simultaneously with any other class.

Antibody polypeptides are encoded by separate variable (V) and constant (C) genes which are subsequently joined at the DNA level during lymphocyte differentiation to code for one polypeptide chain. Considerable chemical, genetic, and serological evidence has accumulated supporting the 'two gene-one polypeptide' theory proposed by Dreyer and Bennett in 1965⁶. Different V regions, as defined by subgroup-specific amino acid sequences, allotypic or idiotypic determinants, are found in association with a single C region for each of the κ and λ light (L) chain and heavy (H) chain loci⁷⁻⁹. Conversely, one V_H region can be shared by several different C_H regions¹⁰⁻¹⁵. In view of the finding by Hozumi and Tonegawa¹⁶ that V and C genes coding for L chains are on separate restriction fragments in embryonic DNA but are on the same restriction fragment in myeloma DNA, it is apparent that gene translocation may have a central role in immunoglobulin synthesis and cell differentiation¹⁷.

As there are multiple C_H genes that can be associated with a single V_H gene, the question arises as to whether the C_H genes are expressed in a sequential order following antigen stimulation. Several studies suggest that one cell or its clonal progeny can produce more than one immunoglobulin class. B cells have been shown to bear membrane receptor molecules of the IgM and IgD class with the same idiotype¹⁸. Using the splenic focus assay, we have previously demonstrated that a single antigen-stimulated B cell can produce clonal progeny making antibodies with different H chain classes that seem to have the same variable region by idiotype analysis¹⁹. Thus, an isolated B-cell precursor specific for the phosphorylcholine (PC) determinant can be stimulated *in vitro* to generate daughter plasma cells which secrete antibody of three immunoglobulin classes, μ , γ 1, and α . Also, antibody molecules of each class have the same idiotype as that of the TEPC 15 (T15) myeloma. As shown here, an analysis of the sequence of appearance of constant regions within such T15 clones suggests that classes of immunoglobulin are produced in a random fashion following antigen stimulation in the presence of carrier help.

Ten million spleen cells from non-immune BALB/c mice were transferred intravenously into lethally irradiated BALB/c recipients that had been primed with 0.1 mg of *Limulus polyphemus* haemocyanin 2 months previously. Sixteen hours after transfer, spleen fragment cultures were prepared from the recipients and stimulated *in vitro* with 5×10^{-7} M 3-(azo-phenylphosphorylcholine)-N-acetyltyrosylglycylglycine haemo-

cyanin²⁰. The antigen was removed by day 4, and culture fluids were collected every 2 d until day 18. Anti-PC antibody was quantified by radioimmunoassay for heavy chain class and idiotype using T15 protein and monoclonal culture fluids as standards. The binding of antibody to antigen coupled to bromoacetylcellulose immunoadsorbant was detected with ¹²⁵I-labelled goat anti-mouse μ , γ 1, and α -chain antibody. Antibody with the T15 idiotype was detected in a radioimmunoassay using anti-T15 sera raised in A/He mice²¹. The monoclonal nature of antibody from these cultures has been demonstrated in previous experiments, indicating that each clone is derived from a single B cell^{19,22}.

The serial analysis of immunoglobulin class production by clones with the T15 idiotype indicates variation in time of appearance after stimulation, amount of antibody, and sequence related to other classes. The overall view of the onset of antibody production, shown in Fig. 1, suggests that any one of the three classes of immunoglobulin may be synthesised initially, and that IgG secretion persists longer than secretion of the other two classes. The apparently random expression of each immunoglobulin class is confirmed by data collected from individual clones which are presented in Table 1. B-cell clones, although making antibody with the same idiotype, display considerable variation in the rate of IgM, IgG, and IgA production with time. Immunoglobulin class expression seems to be completely random in that any class of immunoglobulin can appear at any time, either singly or in combination with others. Variations in time of synthesis after stimulation and in amount of antibody in each immunoglobulin class produced by individual clones are observed. In an analysis of over 60 clones making antibody with the T15 idiotype, the early production of solely IgM and later solely IgG (or IgA) was not observed. Indeed, some clones made IgG before or without detectable IgM synthesis.

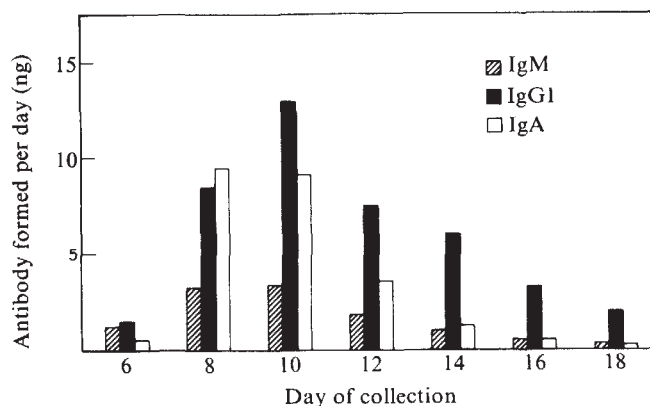


Fig. 1 Heavy chain class secretion by PC-specific clones with the T15 idiotype. The average amount of antibody produced by 43 clones in each immunoglobulin class is plotted as a function of time. Monoclonal fragment cultures were stimulated with antigen on day 1 and the antigen was removed on day 4. Culture fluids were collected every two days after that and analysed for idiotype and C_H determinants by radioimmunoassay.

The idea of an IgM to IgG 'switch' arose from observations that IgM is detected before IgG antibody in immune sera^{1,2}. Several explanations may be offered for these results. First, the assays used for IgM detection are more sensitive than those used for IgG detection. Second, it may be necessary to generate carrier-specific help to allow maximum expression of classes of immunoglobulin other than IgM; that is, IgM may appear first during the immune response because it requires less carrier-specific help. Thus, in physiological conditions, T cells or factors may influence the expression of heavy chain classes. Finally, subsets of B cells may have different developmental potentials for expressing C_H regions. For example, the clonal progeny of immature B cells produce predominantly IgM, whereas clones derived from immune B cells produce predominantly IgG²². Nonetheless, using a method which detects each heavy chain class equally well and provides

Table 1 Random production of immunoglobulin class by B-cell clones with time after antigen stimulation

Individual clones with T15 idiotype	Heavy chain class	Antibody formed on day of collection after antigen stimulation (ng)						
		Day 6	8	10	12	14	16	18
1	μ	0	6	4	1	0	0	0
	$\gamma 1$	0	0	0	0	0	0	0
	α	0	0	0	0	0	0	0
2	μ	0	0	0	0	0	0	0
	$\gamma 1$	1	12	6	2	0	0	0
	α	0	9	0	0	0	0	0
3	μ	0	0	0	0	0	0	0
	$\gamma 1$	0	0	0	0	0	0	0
	α	0	2	16	43	16	0	0
4	μ	0	0	1	2	0	0	0
	$\gamma 1$	2	4	15	10	9	2	3
	α	0	0	0	0	0	0	0
5	μ	0	14	16	5	2	2	0
	$\gamma 1$	0	1	2	0	0	0	0
	α	0	16	34	8	5	1	0
6	μ	2	3	9	3	1	0	0
	$\gamma 1$	11	60	75	103	31	23	8
	α	10	44	24	5	2	1	0
7	μ	7	11	11	8	6	2	1
	$\gamma 1$	1	6	9	7	8	3	2
	α	16	122	185	172	43	23	16
8	μ	0	0	11	2	0	0	0
	$\gamma 1$	0	0	28	18	5	8	6
	α	0	0	0	0	0	0	0
9	μ	2	7	7	2	0	0	0
	$\gamma 1$	1	12	10	3	0	0	0
	α	2	69	39	3	0	0	0
10	μ	1	4	11	5	2	2	2
	$\gamma 1$	0	3	19	34	6	2	0
	α	0	0	2	3	0	0	0

maximal carrier help for the stimulation of nonimmune B cells, the data presented here clearly indicate that the onset of expression of different immunoglobulin classes is completely random in that any class can appear singly or in combination with others. In this regard, no evidence was found for a sequential shift from $\mu \rightarrow \gamma 1 \rightarrow \alpha$ by stimulated clones.

There are two general models for the association of V_H genes with their corresponding C_H genes^{1,5}. The simultaneous insertional model suggests that the V_H gene is copied multiple times and is associated with two or more C_H genes. The successive insertional model proposes that the V_H gene is successively translocated to various C_H genes during differentiation. The data presented here suggest that if the first model is correct, random expression of a V_H - C_H gene combination occurs, whereas if the second model is correct, a V_H gene can be translocated next to any C_H gene following antigen-triggered differentiation. These two models can probably only be tested by direct studies at the nucleic acid level.

This work was supported by USPHS grants from USNIH and NCI to Dr Norman Klinman at the University of Pennsylvania and a Helen Hay Whitney fellowship. I thank N. Klinman for encouragement and L. Hood for criticisms.

PATRICIA J. GEARHART

Department of Biology,
The Johns Hopkins University,
Baltimore, Maryland 21218

Received 27 June; accepted 15 August 1977.

1. Bauer, D. C. & Stavitsky, A. B. *Proc. natn. Acad. Sci. U.S.A.* **47**, 1667-1680 (1961).
2. Nossal, G. J. V., Szenteg, A., Ada, G. L. & Austin, C. J. *exp. Med.* **119**, 485-502 (1964).
3. Papermaster, B. W. *Cold Spring Harb. Symp. quant. Biol.* **32**, 447-460 (1967).
4. Lawton, A. R., Asofsky, R., Hylton, M. B. & Cooper, M. D. *J. exp. Med.* **135**, 277-297 (1972).
5. Herrod, H. G. & Warner, N. L. *J. Immun.* **108**, 1712-1717 (1972).

6. Dreyer, W. J. & Bennett, J. C. *Proc. natn. Acad. Sci. U.S.A.* **54**, 864-869 (1965).
7. Hood, L., Gray, W. R., Sanders, B. G. & Dreyer, W. J. *Cold Spring Harb. Symp. quant. Biol.* **32**, 133-146 (1967).
8. Hilschmann, N. Z. *physiol. Chem.* **349**, 945-951 (1968).
9. Cunningham, B. A., Pflumm, M. N., Rutishauser, U. & Edelman, G. M. *Proc. natn. Acad. Sci. U.S.A.* **64**, 997-1003 (1969).
10. Todd, C. W. *Biochem. biophys. Res. Commun.* **11**, 170-175 (1963).
11. Wikler, M., Köhler, H., Sinoda, T. & Putnam, F. W. *Science* **163**, 75-78 (1969).
12. Oudin, J. & Michel, M. J. *exp. Med.* **130**, 619-642 (1969).
13. Penn, G. M., Kunkel, H. & Grey, H. *Proc. Soc. exp. Biol. Med.* **135**, 660-665 (1970).
14. Wang, A. A., Wilson, S. K., Hopper, J. E., Fudenberg, H. H. & Nisonoff, A. *Proc. natn. Acad. Sci. U.S.A.* **66**, 337-343 (1970).
15. Sledge, C., Fair, D. S., Black, B., Krueger, R. G. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **73**, 923-927 (1976).
16. Hozumi, N. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3628-3632 (1976).
17. Gally, J. A. & Edelman, G. M. *Nature* **227**, 341-348 (1970).
18. Fu, S. M., Winchester, R. J. & Kunkel, H. G. *J. Immun.* **114**, 250-252 (1975).
19. Gearhart, P. J., Sigal, N. H. & Klinman, N. R. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1707-1711 (1975).
20. Sigal, N. H., Gearhart, P. J. & Klinman, N. R. *J. Immun.* **114**, 1354-1358 (1975).
21. Gearhart, P. J., Sigal, N. H. & Klinman, N. R. *J. exp. Med.* **141**, 56-71 (1975).
22. Klinman, N. R. & Press, J. L. *Transplant. Rev.* **24**, 41-83 (1975).

Immunoprotection by embryonal carcinoma cells for methyl-cholanthrene-induced murine sarcomas

METHYLCHOLANTHRENE-induced sarcomas of mice (MCS) usually have transplantation antigens specific to each independently arising tumour¹. These antigens, defined by *in vivo* protection assays, are different from those defined serologically as viral² or embryonic antigens^{3,4}. The latter, common to most sarcomas, seem to contribute little in transplantation assays as cross protection is rare and weak^{1,5}. The occasional success in protecting against sarcoma growth by previous immunisation with cells of mid-term embryos^{4,6} encourages the search for more effective ways of using common embryonic antigens for this purpose. We

show here that effective immunisation against several different non-cross-reacting sarcomas, including one not immunogenic itself, can be achieved with embryonal carcinoma cells. These cells are the stem cells of a teratocarcinoma and show remarkable similarities to the pluripotent cells of early mouse embryos with regard to cell surface antigens⁷ as well as in morphological and biochemical properties⁸.

The sarcomas used were a set induced in female C57BL/10ScSn (B10) mice by subcutaneous injection of 0.5 mg 3-methylcholanthrene into the hind legs. Approximately 100 d later primary tumours were excised, passaged subcutaneously twice and then stored in liquid nitrogen. Tumours (*in vivo* passages 2–8) were characterised for immunogenicity in a rejection assay (Fig. 1). To assay immunogenicity, mice were immunised in two ways: by allowing tumour growth for 10 d followed by tumour excision or by three weekly immunisations with 10^7 irradiated MCS cells (legend Fig. 1). The mice were challenged a week later with 2×10^5 viable cells of the same or a different sarcoma and tumour diameter measured as a function of time.

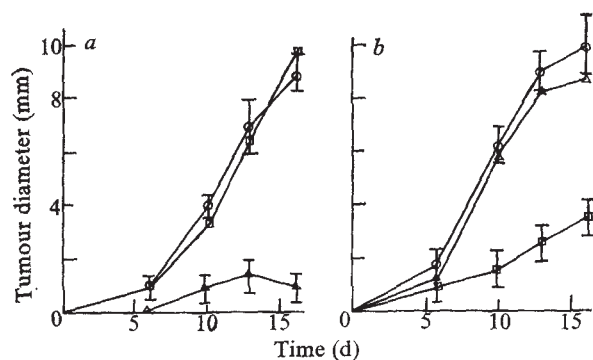


Fig. 1 Immunogenicity of methylcholanthrene-induced sarcomas MC6A and MC6B. Groups of 12 female B10 mice, 12–16 weeks old were immunised by three weekly subcutaneous injections of 10^7 MC6A or MC6B cells (irradiated with 10,000 rad from a ^{60}Co source) in 0.2 ml phosphate-buffered saline (PBS) into the right groin. Tumour cell suspensions were prepared from *in vivo* tumours chopped finely and stirred continuously at 37°C in 2.5% trypsin in Eagles Ca^{2+} - and Mg^{2+} -free medium. One week following the last immunisation, groups of six mice were challenged with 2×10^5 viable MC6A (a) or MC6B (b) cells in 0.2 ml PBS into the right groin. Control groups of unimmunised mice were similarly challenged. Tumour diameter was measured in two perpendicular planes every 3 d using calipers for 16 d after challenge. Error bars represent $\pm$ s.e.m. $\circ$, unimmunised; Δ , MC6A immunised; $\square$, MC6B immunised.

Figure 1 shows that MC6A and MC6B are strongly immunogenic; immunisation evokes good protection against subsequent challenge with the same methylcholanthrene sarcoma. A measure of the strength of immunisation is the antigenic ratio (AR) (ref. 9) (Table 1). The relative immunogenicity measured by AR ranks the tumours in the descending order MC6A, MC6B, MC9, MC5 with MC5 being non-immunogenic (Table 1). Data in Fig. 1 and Table 1 also demonstrate that MC6A and MC6B do not cross protect and thus apparently possess individually specific antigens.

Similar immunisations with cells of an embryonal carcinoma cell line Nulli-SCC.1, which do not differentiate *in vivo* or *in vitro*¹⁰, elicit protection against the strongly immunogenic MC6A and MC6B, the less immunogenic MC9, and the non-immunogenic MC5 (Table 1, Fig. 2a, b and c). The Nulli-SCC.1 cells are derived from the 129/J mouse strain which is not syngeneic with B10. To test whether alloantigens on Nulli were responsible for the protective effect against the sarcomas, 129/J spleen cells were used to immunise B10 mice; no protection was seen (Fig. 2d). In addition, immunisation with cells of C57BL/6 thymoma EL4 (ref. 11), a DBA/2 mastocytoma P815 (ref. 12), or a B10A Abelson lymphoma induced in our laboratory, gave no protection against subsequent challenge with MC6A cells (Fig. 2d). The Nulli-SCC.1 cells grow as homogeneous embryonal carcinoma *in vitro* and the cell line was derived from a spontaneous testicular teratocarcinoma (LS-402C-1684). These Nulli-SCC.1 cells resemble

Table 1 Relative immunogenicity of methylcholanthrene sarcomas and embryonal carcinoma with MCS challenge

Immunising tumour*	Challenge tumour†	Antigenic ratio‡
MC6A	MC6A	10.0
MC6B	MC6A	1.1
MC6A	MC6B	1.0
MC6B	MC6B	5.3
MC9	MC9	3.2
MC5	MC5	1.0
Nulli-SCC.1	MC6A	2.1
Nulli-SCC.1	MC6B	2.0
Nulli-SCC.1	MC9	1.9
Nulli-SCC.1	MC5	1.7

* 10^6 viable tumour cells prepared as in Fig. 1 were injected in 0.2 ml phosphate buffered saline (PBS) into the left flank of groups of four mice and tumours excised surgically 10–14 d later. Nulli-SCC.1 cells were grown as monolayers in tissue culture and cell suspensions prepared as previously described⁸.

† 2×10^5 viable cells were injected subcutaneously (s.c.) in 0.2 ml PBS into the right groin 10 d after the excision. Tumour diameter was measured in two perpendicular planes every 3 d by calipers and the mean calculated. A control group of unimmunised animals was challenged concurrently.

‡Antigenic ratio is calculated as

$$\frac{\text{mean tumour diameter in unimmunised mice}}{\text{mean tumour diameter in immunised mice}}$$

at 16 d after challenge. An antigenic ratio of greater than 1.0 indicates protection. Mean tumour diameter is computed from all mice in a group; no challenge dose failed to grow in these experiments.

early embryo cells in morphological, biochemical and antigenic properties¹³, but differ because they can no longer differentiate. Immunisation with irradiated homogeneous embryonal carcinoma cells from a pluripotent teratocarcinoma cell line¹⁰ also shows protection against challenge with MC6A tumour cells. This line was isolated from a 3-d embryo induced teratocarcinoma (OTT-5568) and retains the ability to differentiate *in vivo* and, in appropriate conditions, *in vitro*.

In addition to embryonal carcinoma, some protection against methylcholanthrene tumour challenge is also provided by immunisation with teratocarcinoma derived endodermal cells of the lines PSA5E (ref. 13) and PYS-2 (ref. 14) (Fig. 2e), though the

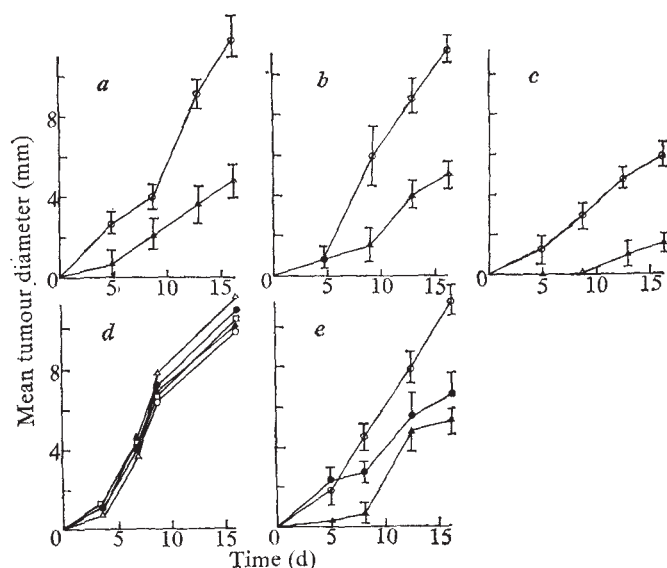


Fig. 2 Protection by various cells against MCS challenge. Groups of 3 B10 female mice were immunised by three weekly s.c. injections of 10^7 cells. A week following the last immunisation, the mice were challenged with 2×10^5 viable tumour cells as in Fig. 1. a, b, and c. Mice immunised with irradiated Nulli-SCC.1 (Δ); unimmunised ($\circ$). Challenge with MC6A, MC6B, MC5 respectively. d. Mice immunised with irradiated EL4 (Δ); B10.A.2.P5 Abelson lymphoma ($\bullet$); P815 mastocytoma ($\blacktriangle$); and 129 spleen cells ($\square$); unimmunised ($\circ$). Challenge with MC6A. e. Mice immunised with PYS ($\bullet$), PSA5E endodermal cells (Δ); unimmunised ($\circ$). Challenge with MC6A.

degree of protection by PYS is weak. These lines were derived from the outgrowths of embryoid bodies *in vitro*.

How do embryonal carcinoma cells induce protection to tumours that do not protect against each other (MC6A and MC6B) and to a tumour that cannot protect against itself (MC5)? Nonspecific activation of the immune system by the embryonic cell lines, though not by the other tumour lines tested, might explain the observed cross protection¹⁵. A more interesting possibility is that there are antigens shared by all these cells and that presentation of these antigens on the embryonal carcinoma cells is particularly effective in inducing specific immune protection while presentation of these common antigens on the MCS is not effectively immunogenic. Indeed serological studies with teratocarcinoma have demonstrated antigen specificities present on these cells and which are shared by a wide range of tumour cell lines^{16,17}. Antisera prepared by us in syngeneic mice against MCS cross react with embryonal carcinoma cells. In addition, lymphocytes sensitised *in vitro* to syngeneic teratocarcinoma cells have been shown to inhibit growth of a wide range of tumour cells *in vitro* and *in vivo*, but do not affect untransformed cells¹⁸.

Why the MCS-embryonal carcinoma shared antigens might be immunogenic only on the latter cells is intriguing. It is perhaps relevant that both the embryonal carcinoma and the endodermal cells do not express the major histocompatibility antigens (H2) (refs 13, 19). Since there is increasing evidence that H2 plays an important part in directing the immune response²⁰, it is possible that in the absence of H2 expression, other antigens, not seen in the presence of H2, become immunodominant.

K.S. holds an MRC Training Fellowship. We thank Susan Baker for technical assistance.

KAROL SIKORA
PETER STERN
EDWIN LENNOX

MRC Laboratory of Molecular Biology,
Hills Road, Cambridge, UK

Received 1 July; accepted 18 August 1977.

- Basombrino, M. A. *Cancer Res.* **30**, 2458-2462 (1970).
- Grant, J. D., Bigner, D. B., Fischinger, P. J. & Bolognesi, D. P. *Proc. natn. Acad. Sci. U.S.A.* **71**, 5037-5041 (1974).
- Brown, R. J. *Int. J. Cancer* **6**, 245-249 (1970).
- Baldwin, R. W., Embleton, M. J., Price, M. R. & Vose, B. M. *Transplant. Rev.* **20**, 77-99 (1974).
- Reiner, J. & Southam, C. M. *Cancer Res.* **27**, 1243-1247 (1967).
- Coggins, J. H. & Anderson, N. G. *Adv. Cancer Res.* **19**, 105-165 (1974).
- Erickson, R. P. in *Immunobiology of Tumors* (eds Johnson, M. & Edidin, M.) 85-114 (Cambridge University Press, Cambridge, 1977).
- Martin, G. R. *Cell* **5**, 229-243 (1975).
- Bartlett, G. L. *J. natn. Cancer Inst.* **49**, 493-504 (1972).
- Martin, G. R. & Evans, M. J. in *Teratomas and Differentiation* (eds Sherman, M. I. & Solter, D.) 169-187 (Academic, London, 1975).
- Gorer, P. A. & Amos, D. B. *Cancer Res.* **16**, 338-343 (1956).
- Dunn, T. B. & Potter, M. J. *J. natn. Cancer Inst.* **18**, 587-595 (1957).
- Evans, M. J., Lovell-Badge, R. & Stern, P. L. in *European Teratology Soc. Meet. Symp.* (in the press).
- Lehman, J. M., Speers, W. C., Swartzendruber, P. E. & Pierce, G. B. *J. Cell Physiol.* **84**, 13-28 (1974).
- Smith, S. E. & Scott, M. T. *Br. J. Cancer* **26**, 361-367 (1972).
- Gooding, L. R. & Edidin, M. *J. exp. Med.* **140**, 61-78 (1974).
- Gooding, L. R., Hsu, Y. C. & Edidin, M. *Dev. Biol.* **49**, 479-486 (1976).
- Bartlett, P. F., Fenderson, B. A. & Edidin, M. *Folia Proc.* **36**, 1290 (1977).
- Arutz, K. & Jacob, F. *Transplantation* **17**, 632-634 (1974).
- Doherty, P. C., Blanden, R. V. & Zinkernagel, R. M. *Transplant. Rev.* **29**, 89-124 (1976).

Chemical mutagenesis of mammalian cells can be quantified

STUDY of mammalian genetics has been traditionally limited to conventional breeding experiments, which are limited by high cost, long duration and the availability of well-characterised genetic markers. These factors have especially impeded rapid progress in mammalian mutagenesis studies. Many of the mechanisms underlying the process of mutation induction in mammals can be investigated through the use of mammalian cells in culture¹⁻⁵, but the use of cultured cells for quantitative analyses of mutagenesis has been hampered by the lack of a suitable mutation-induction system, particularly with chemical mutagens. We show here that quantitative dosimetry of chemical mutagenesis can be studied in a mammalian cell culture system. This will allow

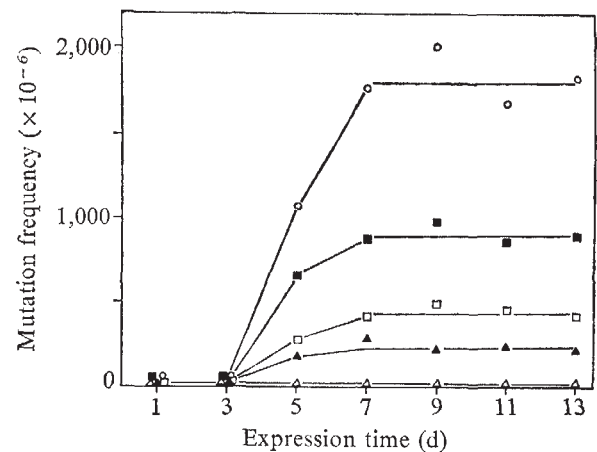


Fig. 1 Expression time of HGPRT⁻ (TG-resistant) mutant phenotype. CHO cells (clone K₁-BH₄) were plated at 5×10^5 cells per 100-mm plate in Ham's F12 medium containing 5% dialysed foetal calf serum and were incubated at 37 °C in 5% CO₂ in air in a 100% humidified incubator. After 24 h (cell no. $1-1.5 \times 10^6$ cells per plate), EMS was added (zero time), and the cultures were incubated for 16 h. The cultures were washed three times with saline, the cells were removed by trypsinisation, and approximately 10^6 cells were subcultured. For selection of the mutant phenotype, cells were plated at 2×10^5 cells per 100-mm plate (5 plates = 10^6 cells) in hypoxanthine-free F12 medium containing 5% dialysed foetal calf serum and 10 µM TG. For cloning efficiency, 200-1,000 cells were plated in triplicate in the same medium without TG. After 7 d of growth colonies were fixed, stained, and counted. Cultures were also subcultured and plated for mutant selection every 2 d thereafter. Mutation frequency was calculated as the number of mutant colonies per 10^6 cells (corrected by cloning efficiency at time of selection). Cells treated with 0 (Δ), 50 (◻), 100 (○), 200 (◼), or 400 (◊) µg ml⁻¹ of EMS. Spontaneous frequency in the range $2-3.6 \times 10^{-6}$.

not only studies of the mechanisms of chemical mutagenesis, but also an estimate of the genetic hazard of chemical agents in the environment.

The mutation assay system we have developed uses cultured Chinese hamster ovary (CHO) cells and the purine analogue 6-thioguanine (TG) to quantitate mutation induction at the hypoxanthine-guanine phosphoribosyl transferase (HGPRT) locus⁶. This CHO/HGPRT system is based on the facts that TG is converted to the nucleotide by the HGPRT activity of wild-type cells and that the nucleotide or its metabolites are lethal. Mutant cells lacking this enzyme activity are able to survive and grow in the presence of TG. We have defined optimal conditions for the selection of the mutant phenotype, including phenotypic expression time, concentration of TG and medium composition for the selection step, and optimal cell density in the selection to permit maximum mutant recovery. With our system the TG-resistant mutation seems to affect the HGPRT locus almost exclusively, since more than 98% of the resistant clones show greatly reduced incorporation of radiolabelled hypoxanthine into macromolecules, are sensitive to the anti-metabolite aminopterin, and contain highly reduced or undetectable HGPRT activity (ref. 6 and O'Neill, J. P., Brimer, P. A., Machanoff, R., Hirsch, G. P., and Hsie, A. W., submitted for publication). Using the CHO/HGPRT system, we have studied mutagenesis with a common chemical mutagen, ethyl methanesulphonate (EMS), in a quantitative fashion.

An important aspect to consider in any mutation induction is phenotypic delay, the time required for mutants to express the altered phenotype. For TG resistance, a delay would be expected due to the time necessary for the dilution and decay of pre-existing HGPRT enzyme and its mRNA before the mutant phenotype could be expressed. This important parameter was investigated with cultures incubated with various concentrations of EMS for 16 h;

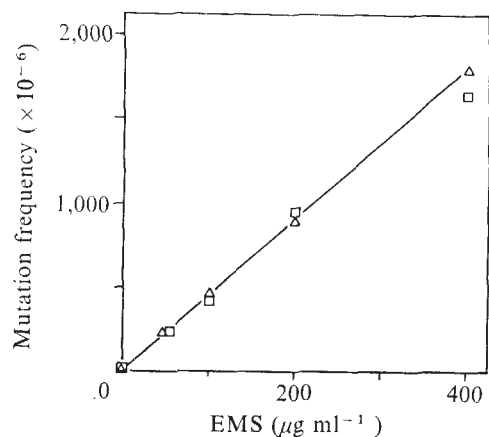


Fig. 2 Dose response of EMS-induced mutagenesis. Data from two separate experiments are presented: the average of the mutation frequencies at days 7, 9, 11, and 13 from Fig. 1 (Δ) (average spontaneous frequency 3.1×10^{-6}); and the frequencies for 24-h incubation time from Fig. 3 ($\square$).

thereafter, the cells were subcultured and plated for selection every 2 d. Mutant colonies were found after 5 d of expression, and the mutation frequency reached a maximum value at day 7 with all EMS doses (Fig. 1). The expression time was independent of the mutagen dose, and there was no loss of mutants with longer periods of expression, as the mutation frequency remained at a constant value in cells subcultured for an additional 6 d. There was a linear dose-response relationship between mutation frequency and EMS concentration (Fig. 2). Since the mutation frequency reached a stable maximum value at 7–9 d of expression, cells were routinely plated for selection on day 9 in all subsequent experiments.

The dose-response data of Fig. 2 were obtained with long incubation times with EMS (16–24 h). It would be difficult to compare such chemical mutagenesis with that induced by physical agents like X rays because of complicating factors such as the stability of EMS in the medium, rate of cellular uptake, and intracellular stability. To investigate these aspects, cultures were incubated with EMS for

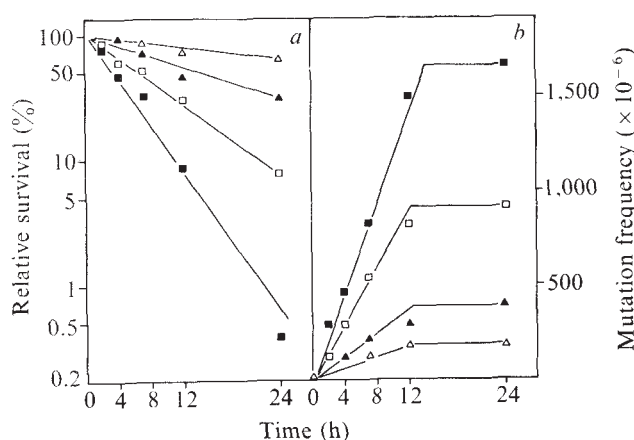


Fig. 3 Initial cell survival (a) and EMS-induced mutation frequency (b) as a function of incubation time. Cells were plated at 5×10^5 cells per plate as described in Fig. 2. After 24 h, EMS was added, cultures were incubated for the indicated time interval and then washed, and fresh medium was added. After a total interval of 24 h the cultures were trypsinised, and single-cell survival was determined (absolute cloning efficiency of untreated control cultures was 84%). Cultures were subcultured every 2 d and plated for mutant selection on day 9 of expression as described in Fig. 1. EMS at concentrations of 50 (Δ), 100 ($\blacktriangle$), 200 ($\square$), and 400 $\mu\text{g ml}^{-1}$ ($\blacksquare$). Spontaneous mutation frequency 4.2×10^{-6} .

intervals of 2–24 h. A linear increase in mutation frequency with incubation times of up to 12–14 h was found with EMS concentrations of 50–400 $\mu\text{g ml}^{-1}$ (Fig. 3b). But cell survival decreased exponentially with time over the entire 24-h period (Fig. 3a). This difference in the time courses of cellular lethality and mutagenicity might be due to the formation of toxic, non-mutagenic breakdown products in the medium with longer incubation times or might reflect a difference in the mode of action of EMS in these two biological effects. The increases in both cell lethality and induced mutation with incubation times of up to 12 h allow a study of the dosimetry of this chemical mutagen. For better defined dosimetry studies, cultures were incubated with various concentrations of EMS (0.05–3.2 mg ml^{-1}) for 2–12 h. As shown in Fig. 4 the decrease in cell survival was exponential with dose ($\text{mg ml}^{-1} \text{ h}$), and the increase in mutation frequency was linear with dose. From these studies the mutagenic potential of EMS can be described as 310×10^{-6} mutants ($\text{cell mg ml}^{-1} \text{ h}$) $^{-1}$.

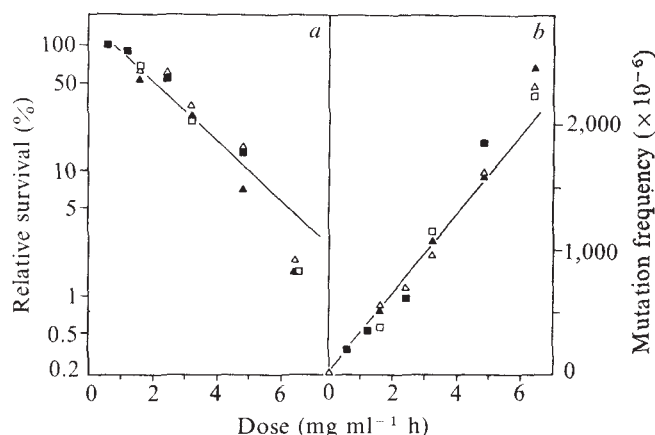


Fig. 4 Dosimetry for EMS cellular lethality (a) and induced mutation frequency (b). Cells were treated with various concentrations (mg ml^{-1}) of EMS for 2 (Δ), 4 ($\blacktriangle$), 8 ($\square$), or 12 h ($\blacksquare$). After treatment the cultures were washed, and single-cell survival was determined after a total time of 24 h. Cells were plated for selection on day 9 of expression as in Fig. 2. Spontaneous mutation frequency = 2.3×10^{-6} . In (a) the linear fit of the survival curve was found over 1.6–6.4 $\text{mg (of EMS) ml}^{-1} \text{ h}$ through fitting to the 'multitarget model'

$$R_t = 1 - (1 - e^{-Bx})^N$$

where B and N are fitted coefficients. The extrapolation number (N) is 1.74 with 95% confidence limits of 1.18, 2.58; the slope (B) is -0.248 with 95% confidence limits of -0.308 , -0.188 . In (b) the straight-line fit for the pooled data on mutation induction over the dose range 0.64–6.4 $\text{mg (of EMS) ml}^{-1} \text{ h}$ is $f(x) = 10^{-6} (5.58 + 310.77x)$. The 95% confidence limits for the slope are 283.68, 337.87. All confidence limits are estimated based on the residual sum of squares. Statistical analyses were performed as in ref. 6.

These results demonstrate (1) that the mutagenicity of EMS in mammalian cells follows a linear dosimetry relationship, and (2) that a quantitative study of chemical mutagenesis in mammalian cells is possible if an adequately defined selection protocol is developed.

We thank M. H. Hsieh, D. G. Gosslee and T. J. Mitchell for help with statistical analysis and D. B. Couch and E. I. Shaw for reviewing the manuscript. This work was supported by the US National Center for Toxicological Research, EPA and ERDA under contract with the Union Carbide Corporation.

J. P. O'NEILL
A. W. HSIEH

Biology Division,
Oak Ridge National Laboratory,
Oak Ridge, Tennessee 37830

Received 27 March; accepted 18 July 1977.

- ¹ Puck, T. T. *The Mammalian Cell as a Microorganism* (Holden-Day, San Francisco, 1972).
² Thompson, L. H. & Baker, R. M. in *Methods in Cell Biology* (ed. Prescott, D. B.) 209–281 (Academic, New York, 1973).
³ Chu, E. H. Y. *Genetics* 78, 115–132 (1974).
⁴ Siminovich, L. *Cell* 7, 1–11 (1976).
⁵ DeMars, R. *Mutat. Res.* 24, 335–364 (1974).
⁶ Hsie, A. W., Brimer, P. A., Mitchell, T. J. & Gosslee, D. G. *Somat. Cell Genet.* 1, 247–261 (1975); 1, 383–389 (1975).

Increasing ovulation rate in sheep by active immunisation against an ovarian steroid androstenedione

ALTHOUGH low rates of ovulation constitute a major limitation to increased productivity of sheep and cattle little is known about the mechanisms that control ovulation either within or among species. Androstenedione, a weak androgen secreted by ovaries¹ and adrenals¹, is quantitatively the major steroid secreted by the ovary before ovulation in several species² and as such it might have a role in determining ovulation rate; it is also an important precursor for the biosynthesis of oestrogens by the ovary³. Although androstenedione can act as a prehormone by virtue of its peripheral metabolism to oestrogens⁴ a direct hormonal action has only been demonstrated in the female rhesus monkeys⁵ where it seems to be important for sexual behaviour. We report here that by actively immunising sheep against androstenedione a slight but significant increase in ovulation rate can be induced.

Two series of experiments were carried out on 10 Welsh mountain ewes. Half of the sheep were immunised against androstenedione conjugated in the 11 position to bovine serum albumin (BSA)⁶ and the other half were immunised against BSA alone and served as controls. All of the ewes immunised against androstenedione-11-BSA produced antibodies with a high degree of specificity for androstenedione⁶, and the titre of antibody was maintained at elevated levels by periodic booster immunisations⁶.

During the normal sheep mating season (October–March) the ewes were kept with a harnessed vasectomised ram and observed daily for signs of oestrus⁷. In March 1975, and again in November 1975, the ewes were subjected to a laparotomy and the number of corpora lutea and visible Graafian follicles of greater than 4 mm diameter were recorded. The pooled results are presented in Table 1 and show that there was a significant increase in the ovulation rate (Fisher's exact test, $P < 0.01$, February 1975; $P < 0.07$, November 1975) but the length of the oestrous cycle and the duration of oestrus were unaffected (t test). The increase in ovulation rate was accompanied by an increase in follicular growth, the control ewes having 1.6 ± 0.2 Graafian follicles per animal in contrast to 4.2 ± 0.7 per ewe in the immunised group ($P < 0.05$, Wilcoxon rank sum test).

Table 1 Length of oestrous cycle, duration of oestrus and ovulation rate in sheep immunised against androstenedione-11-BSA

Group	Oestrous cycle (d) (mean $\pm$ s.e.m.)	Duration of oestrus (h) (mean $\pm$ s.e.m.)	Ovulation rate (mean $\pm$ s.e.m.)
Control	16.7 ± 0.2 ($n = 56$)	40.4 ± 1.3 ($n = 10$)	1.11 ± 0.11 ($n = 9$)
Immunised	17.5 ± 0.8 ($n = 70$)	45.1 ± 1.9 ($n = 7$)	2.13 ± 0.12 ($n = 8$)

These experiments clearly show that ovulation rate can be increased by immunisation against androstenedione-11-BSA. This technique has the additional advantage that excessive superovulation such as is often produced by the use of exogenous gonadotrophins is avoided; all of the immunised ewes had either two or three ovulations, while all the control ewes had one or two ovulations. In subsequent experiments on larger groups of sheep only 1 out of 15 immunised ewes had an ovulation rate of greater than 3.

Measurement of the peripheral levels of luteinising hormone

(LH) (ref. 8) and progesterone⁶ in samples collected daily throughout one complete oestrous cycle are shown in Fig. 1. The ewes immunised against androstenedione-11-BSA had elevated peripheral concentrations of both hormones throughout most of the oestrous cycle. An additional series of blood samples was collected from both groups of ewes, the collections beginning at 0800 LT on day 16 of the oestrous cycle and continuing at 4-h intervals until 36 h after the onset of oestrus. The plasma samples were assayed for the LH content, and the results show that the preovulatory LH discharge reached its maximum concentration 24 ± 4.0 h (mean $\pm$ s.e.m. $n = 3$) after the onset of oestrus in the immunised ewes as compared with 8 ± 1.3 h ($n = 5$) in the control group ($P < 0.001$; unpaired t test).

These experiments suggest that androstenedione has a role (possibly direct) in the control of ovulation rate and, furthermore, that androstenedione might provide a manifestation of the genetic expression of ovulation rate. The mechanisms involved in

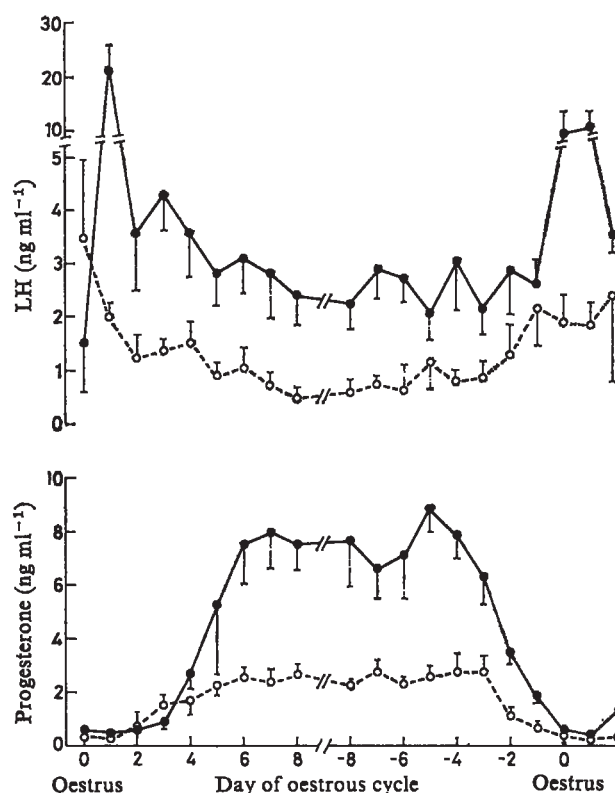


Fig 1 The mean $\pm$ s.e.m. peripheral concentrations of progesterone (lower) and LH (upper) throughout the oestrous cycle of ewes immunised against androstenedione-11-BSA ($\bullet$, $n = 5$) and control ewes immunised against BSA ($\circ$, $n = 5$). The oestrous cycle is dated 8 d forward and 8 d back from successive oestrous periods. Note the split vertical scale for LH.

producing the increased ovulation rate and peripheral concentrations of progesterone and LH are speculative. Neutralisation of the biological activity of androstenedione by antibody binding could interfere with oestrogen synthesis in the ovary, or with the conversion of androstenedione to oestrogen in peripheral tissues and hypothalamus. Alternatively, androstenedione itself may have a direct inhibitory action on the hypothalamo-hypophyseal system. It is even possible that the androstenedione antibody cross reacts to a minor degree with oestrogens, although this cross reaction was shown to be less than 0.1% in an *in vitro* test⁸. Whatever the explanation, our results show that the ovarian steroid hormone(s) involved in the feedback control of LH have a reduced effectiveness after immunisation against androstenedione-11-BSA, and that this exposes the animal to higher tonic secretions of gonadotrophin which in turn leads to increased follicular development, multiple ovulations and enhanced progesterone secretion. A similar mechanism is thought

to operate when synthetic anti-oestrogens such as clomiphene are used to induce ovulation in anovulatory women⁹. The ability to produce a modest increase in the ovulation rate in sheep following a simple immunisation procedure could be of practical significance, provided that the increased ovulation rate leads to an increased number of live births per sheep mated.

We thank Dr G. Woods (Organon Laboratories, Newhouse, Scotland) and Dr F. Rutherford MRC Radioimmunoassay Team, Edinburgh) for preparing the androstenedione-11-BSA, and Mr J. B. Donnelly, CSIRO Division Mathematics and Statistics, for statistical advice. P. Van Look was supported by the National Fonds voor Wetenschappelijk Onderzoek, Brussels, Belgium.

R. J. SCARAMUZZI*
W. G. DAVIDSON
P. F. A. VAN LOOK†

MRC Unit of Reproductive Biology,
2 Forrest Road,
Edinburgh, UK

Received 11 January; accepted 9 August 1977.

*Present address: C.S.I.R.O., Division of Animal Production, P.O. Box 239, Blacktown, N.S.W. 2148, Australia.

†Present address: Vrouwenklinik, Academisch Ziekenhuis Leiden, Leiden, The Netherlands.

1. Baird, D. T., McCracken, J. A. & Goding, J. R. in *Pregnancy and Parturition* (ed. Pierpoint, C. G.) 5–21 (Alpha Omega, Cardiff, 1973).
2. Baird, D. T. in *The Ovary* 2nd edn (eds Zuckerman, Sir S. & Weir, B.) (Academic, New York, London, in the press).
3. Rado, A., McCracken, J. A. & Baird, D. T. *Acta endocr., Copenh.* **65**, 244–260 (1970).
4. Baird, D. T., Horton, R., Longcope, C. & Tait, J. F. *Perspect. Biol. Med.* **11**, 384–421 (1968).
5. Everitt, B. J. & Herbert, J. J. *Endocr.* **51**, 575–588 (1971).
6. Scaramuzzi, R. J., Corker, C. S., Young, G. & Baird, D. T. in *Steroid Immunoassay* (eds Cameron, E. H. P., Hillier, S. G. & Griffith, K.) 111–122 (Alpha Omega, Cardiff, 1975).
7. Radford, H. M., Watson, R. H. & Wood, G. F. *Aust. vet. J.* **36**, 57–66 (1960).
8. Martensz, N. D., Baird, D. T., Scaramuzzi, R. J. & Van Look, P. F. A. *J. Endocr.* **69**, 227–237 (1976).
9. Vaitukaitis, J. L., Bermudez, J. A., Cargille, C. M., Lipsett, M. B. & Ross, G. T. *J. clin. Endocr. Metab.* **32**, 503–508 (1971).

Is α MSH a trophic hormone to adrenal function in the foetus?

FUNCTIONAL development of the pituitary–adrenal axis in the foetus seems to be of considerable importance in influencing the maturation of other foetal organ systems, and the timing of parturition^{1,2}. In several species including man^{3,4}, sheep⁵ and rabbits⁶, there is evidence for an increase in foetal adrenal function in late pregnancy, which seems to depend on the activities of the foetal pituitary and hypothalamus⁷. The factors responsible for stimulating foetal adrenal development during late gestation remain poorly understood, however. From detailed studies in foetal sheep, it is apparent that the prepartum acceleration of adrenal growth and glucocorticoid secretion is not preceded by elevated ACTH concentrations in the foetal plasma^{8,9}. But, prostaglandin E₂ (PGE₂) administered intra-arterially into the foetus stimulated an increase in the concentration of cortisol in foetal plasma⁸ at a stage of pregnancy when the adrenal gland was essentially unresponsive to endogenous or exogenous ACTH (refs 2, 9). Experiments in newborn lambs suggest that the pituitary or brain rather than the foetal adrenal gland was the likely site of the PGE₂ action¹⁰, raising the possibility that other pituitary peptides might mediate this effect. The ACTH related peptides, α -melanocyte stimulating hormone (α MSH) and corticotropin-like intermediate lobe peptide (CLIP) have been identified in human foetal pituitaries, and it has been suggested that the ratio of these peptides to ACTH might be critical to maturation of foetal adrenal function¹¹. We have examined the ability of one of these peptides, α MSH, to influence adrenal function in the rabbit foetus. We present here evidence that the foetal adrenal secretes cortisol in response to α MSH at stages in pregnancy when it is unresponsive to ACTH, but as the sensitivity to ACTH increases, the relative response to α MSH is diminished.

Pregnant New Zealand white rabbits were purchased from Canadian Hybrid Farms, Quebec on day 17 of gestation and were housed individually. The animals had been mated once with a fertile buck, and the day of mating was designated as day 0. Experiments were performed on foetuses on day 22 post coitum and day 28 post coitum and on newborn rabbits, 4–6 d-old after spontaneous delivery at full term.

Foetal experiments were performed under local (1% Xylocaine) anaesthesia. A laparotomy incision was made, the uterus exposed, and ACTH (10 μ g Synacthen, ACTH_{1–24}, CIBA), α MSH (10 μ g, Bachem) or saline was injected intramuscularly (i.m.) into the foetus through the uterine wall. All injections were made in 100 μ l saline. After injection, the uterus was replaced, and the abdomen closed with surgical clips. After 60 min the foetuses were delivered, decapitated, and the blood collected into chilled heparinised tubes and immediately frozen. Newborn rabbits, handled gently with gloves, were injected i.m. with ACTH, α MSH or saline as above, and returned to their mothers. After 60 min the newborns were removed and blood (0.5–1.0 ml) collected rapidly by cardiac puncture. All samples were frozen at -20°C until analysis for cortisol using a previously described antibody⁸ and radioimmunoassay procedure^{8,12}.

An initial study in newborn rabbits examined the time course of the blood cortisol response to ACTH injection. The concentration of cortisol increased from 2.97 ng ml⁻¹ at the time of injection to a maximum concentration (10.9 ng ml⁻¹) within 60 min (Table 1). This time interval was therefore adopted in subsequent studies.

Table 1 Blood cortisol concentrations in foetal and newborn rabbits in response to exogenous α MSH or ACTH

Age	Control t_0	Cortisol (ng ml ⁻¹)		
		Saline t_{60}	Synacthen t_{60}	α MSH t_{60}
22 d p.c.	14.65 $\pm$ 1.02	18.35 $\pm$ 2.85	21.39 $\pm$ 3.24	32.45 $\pm$ 3.17**
28 d p.c.	6.60 $\pm$ 0.82	6.13 $\pm$ 1.57	7.94 $\pm$ 0.83	10.54 $\pm$ 0.80*
4–6 d p.p.	2.97 $\pm$ 1.05	2.74 $\pm$ 0.76	10.91 $\pm$ 1.45***	6.93 $\pm$ 1.36*

The values are the mean $\pm$ s.e. of 10–18 foetuses or newborn from 2–13 mothers in each group. Control observations (t_0) were made before injections. Measurements were then made 60 mins (t_{60}) after saline, Synacthen, or α MSH. p.c., Post coitum; p.p., post partum. Significant differences from 'saline t_{60} ' are indicated as follows: *** P < 0.001, ** P < 0.005, * P < 0.05. Other mean values within any age group are not statistically significantly different from the corresponding saline t_{60} group.

There was no significant effect (P > 0.05) of saline injection on the blood cortisol concentration in any of the three groups of animals (Table 1). In foetuses 22 and 28 d post-coitum, Synacthen produced no significant change in the blood cortisol concentration, whereas a significant increase in cortisol (P < 0.001) was found 60 min after Synacthen injection in the newborn. In marked contrast, α MSH caused a significant increase in the concentration of cortisol in the blood of foetuses at 22 and 28 d post-coitum (Table 1). α MSH also increased the concentration of cortisol in the blood of newborn rabbits but the response relative to that achieved with Synacthen was less marked than earlier in gestation.

Although cortisol and corticosterone are present in similar concentrations in the blood of foetal rabbits³, cortisol binds more strongly than corticosterone to liver, kidney and lung receptor sites¹³, and in adults its secretion is stimulated preferentially by ACTH (ref. 14). We therefore measured cortisol as an index of foetal adrenal glucocorticoid production. The control (t_0) levels we have measured are similar to those reported by Mulay *et al.*⁵, who also observed a decrease in the plasma cortisol concentration between day 22 and 28. In view of the significant increase in body weight and blood volume of the rabbit

foetus between day 22 and 28, comparison of blood cortisol levels between these times are difficult to interpret, and a twofold concentration decrease may not necessarily reflect a decrease in total production rate. Total and unbound cortisol in plasma increase significantly by day 30 (ref. 5), providing evidence for a pre-partum increase in adrenal steroid secretion.

At days 22 and 28 post coitum there was no increase in cortisol in response to ACTH injection into the foetus. Albano *et al.*¹⁵, failed to stimulate adrenal cyclic AMP with ACTH at day 22, implying that the appropriate tropic hormone receptors are poorly developed at that time. A similar insensitivity of the foetal adrenal to ACTH has been reported in the sheep^{2,9} and rhesus monkey¹⁶. It is of interest that while ACTH did not produce short-term changes in the concentration of cortisol in foetal blood, it will maintain adrenal weights in hypophysectomised rabbit foetuses¹⁷, showing that any deficiency of ACTH receptor is not absolute. In contrast to the effects in the foetus, we have shown that in the newborn rabbit, as in the newborn primate¹⁸, ACTH will stimulate an increase in the blood concentration of cortisol.

In striking contrast to the ineffectiveness of ACTH, α MSH provoked a significant increase in the concentration of cortisol in the blood of rabbit foetuses at day 22. The response to α MSH, both absolute and proportional was less marked at day 28, although this could relate to the dose injected. In the newborn the response to α MSH relative to that with Synacthen was reduced further. Previous work has shown that adrenal cells from adult rabbits *in vitro* do not respond to hyperphysiological amounts of α MSH, although they do secrete cortisol in response to ACTH¹⁸. The possible role of hormones other than ACTH as trophic agents to the adrenal gland in the foetus has been suggested from studies in sheep^{8,9,10} and in man^{11,19}, where the ACTH-related peptides, α MSH and CLIP have been isolated from foetal pituitaries. Although the concentrations of these peptides relative to ACTH decreased during late gestation and after birth, it is apparent that they are the dominant peptides of foetal life^{8,11}.

While one cannot conclude from our study that α MSH is 'the' trophic hormone to foetal adrenal function in the rabbit, it is apparent that the foetal adrenal responds to it by day 22 of pregnancy, at a time it does not respond to ACTH. The observation that the foetus may respond to a trophic hormone which is ineffective in the adult is an interesting one, and suggests that it may be inappropriate to extrapolate from adult control mechanisms to those in the foetus. In that sense, it may be necessary to revise some of our concepts of underlying control mechanisms in foetal endocrinology.

We thank Drs F. Naftolin and D. K. Lee for helpful comments. The work was supported by the Fraser Memorial Trust of the Royal Victoria Hospital, the Quebec Research Council and the MRC of Canada. J.R.G.C. is a Canada MRC Scholar.

JOHN R. G. CHALLIS*
JAMES D. TOROSIS

McGill University,
Departments of Obstetrics and Gynaecology
and Physiology,
Montreal, Quebec, Canada

Received 26 July; accepted 26 August 1977.

*To whom correspondence should be addressed. Present address: Royal Victoria Hospital, Women's Pavilion, 687 Pine Avenue West, Montreal, Quebec, Canada.

1. Avery, M. E. *Br. med. Bull.* **31**, 13–18 (1975).
2. Thornburn, G. D., Challis, J. R. G. & Robinson, J. S. in *Biology of the Uterus* (ed. Wynn, R. M.) 653–732 (Plenum, New York, 1977).
3. Murphy, B. E. P., Patrick, J. E. & Denton, R. L. *J. clin. endocr. Metab.* **40**, 164–167 (1975).
4. Fencl, M. de M. & Tulchinsky, D. *New Engl. J. Med.* **292**, 133–136 (1975).

5. Mulay, S., Giannopoulos, G. & Solomon, S. *Endocrinology* **93**, 1342–1348 (1973).
6. Rees, L. H., Jack, P. M. B., Thomas, A. L., Nathanielsz, P. W. *Nature* **253**, 274–275 (1975).
7. Jones, C. T., Boddy, K. & Robinson, J. S. *J. Endocr.* **72**, 293–300 (1977).
8. Louis, T. M., Challis, J. R. G., Robinson, J. S. & Thornburn, G. D. *Nature* **264**, 797–798 (1976).
9. Jones, C. T., Boddy, K., Robinson, J. S. & Ratcliffe, J. G. *J. Endocr.* **72**, 279–292 (1977).
10. Challis, J. R. G., Carson, G. D. & Naftolin, F. *J. Endocr.* (in the press).
11. Silman, R. E., Chard, T., Lowry, P. J., Smith, I. & Young, I. M. *Nature* **260**, 716–718 (1976).
12. Challis, J. R. G., Davies, I. J., Benirschke, K., Hendrickx, A. G. & Ryan, K. J. *Endocrinology* **96**, 185–192 (1975).
13. Giannopoulos, G., Hassan, Z. & Solomon, S. *J. biol. Chem.* **249**, 2424–2427 (1974).
14. Ganjam, V. K., Campwell, A. L. & Murphy, B. E. P. *Endocrinology* **91**, 607–613, (1972).
15. Albano, J. D. M., Jack, P. M., Joseph, T., Gould, R. P., Nathanielsz, P. W. & Brown, B. J. *J. Endocr.* **71**, 333–341 (1976).
16. Jaffe, R. B., Seron-Ferre, M., Huhtaniemi, I. & Korenbrot, C. *J. Steroid Biochem.* (in the press).
17. Jost, A. in *Foetal Autonomy* (eds. Wolstenholme, G. E. W. & O'Connor, M.). *Ciba Fdn Symp.* 79–89 (Churchill, London, 1969).
18. Lowry, P. J., McMartin, C. & Peters, J. *J. Endocr.* **59**, 43–55 (1973).
19. Chard, T., Silman, R. E. & Rees, L. H. in *The Foetus and Birth* (eds. Knight, J. & O'Connor, M.). *Ciba Fdn Symp.* 354–370 (Elsevier, Amsterdam, 1977).

Local anaesthetic benzyl alcohol increases membrane thickness

WE present here evidence that the local anaesthetic, benzyl alcohol, causes a large expansion in the thickness (an increase of about 25% at 7.5 mM concentration) of the hydrocarbon region of bimolecular lipid membranes. Such a change in the bilayer fabric of nerve membranes would provide a new insight into the molecular basis of anaesthesia. While the primary process involved in the mode of action of the anaesthetic is apparently different from the possible effect on the boundary lipids around the sodium channel, recently suggested by Lee¹, the main features remain the same; the sodium channel becomes inoperative (closes?) due to a stress resulting from changes in the membrane's bilayer structure in the presence of the anaesthetic.

The effect of the benzyl alcohol on the thickness of the artificial lecithin bilayer membranes was studied using four-terminal measurements of the very low frequency (0.1–100 Hz) capacitance and conductance of artificial lecithin bilayers with the aid of a digital impedance measuring technique elsewhere^{2–4}.

The effect of benzyl alcohol on the hydrocarbon region capacitance is shown in Fig. 1a. At concentrations of over 7.5 mM, benzyl alcohol decreased the hydrocarbon region capacitance by more than 25%. The conductance of this region also increased in the presence of the anaesthetic (Fig. 1b).

In the absence of the anaesthetic, with a dielectric constant $\epsilon_H=2.1$ for the hydrocarbon region², the bilayer has a hydrocarbon region thickness $\delta_H=3.7$ nm. This is considerably less than twice the fully extended length of the hydrocarbon tails of the predominant lecithins in egg phosphatidylcholine. In the presence of the anaesthetic at 7.5 mM concentration the capacitance data yields a thickness^{2,4} of 4.9 nm for the hydrocarbon region. This is a little larger than twice the fully extended length of the hydrocarbon tails. In the absence of the anaesthetic the hydrocarbon tails therefore seem to be much folded, coiled, kinked (double-gauche forms may predominate) or even somewhat interdigitated.

The anaesthetic apparently has the effect of straightening out the hydrocarbon tails. This process is accompanied by an increase in the conductance of the membranes and, from independent nuclear magnetic resonance studies⁵, an increase in fluidity.

Although the presence of cholesterol in the membrane was found itself to decrease the hydrocarbon region thickness, we found that the effect of benzyl alcohol on lecithin-cholesterol membranes was similar to that reported here for pure lecithin bilayers when the electrolyte was 1 mM KCl.

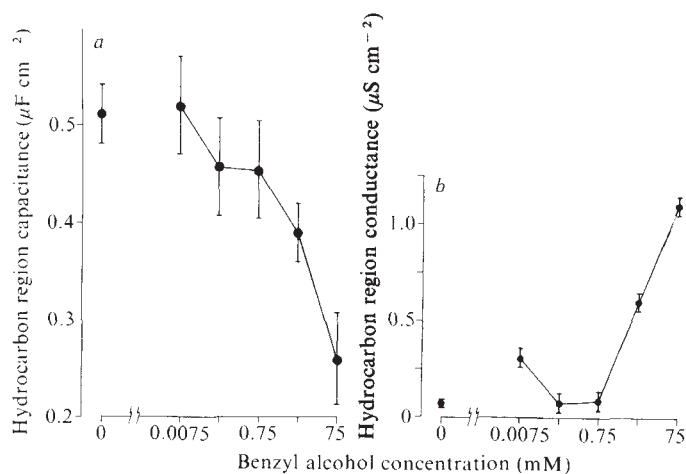


Fig. 1 *a*, The effect of benzyl alcohol on the capacitance of the hydrocarbon layer of bimolecular lipid (lecithin) membranes. The membranes were formed in a 1 mM KCl solution to which the anaesthetic was added. *b*, The conductance of the hydrocarbon region for various concentrations of the anaesthetic. The general increase in the conductance of this region was associated with an increase also in the apparent, mechanical, fluidity of the membranes. The points shown are mean values from several sets of measurements on different membranes (average of 12 measurements per point). The vertical bars indicate the total experimental scatter.

The lecithin bilayer membranes were generated from films of lecithin in *n*-tetradecane and at high concentrations of benzyl alcohol it seems from the low values of the hydrocarbon region capacitance that significant quantities of the *n*-tetradecane may then have been retained in the membrane.

The change in the thickness on addition of the benzyl alcohol at 7.5 mM concentration amounts to ~1.2 nm. This is probably > 15% of the total length of the excitation conduction modules for sodium ions, which span the membrane. Such a strain could readily lead to a closure of the channel. This is shown schematically in Fig. 2. For illustration purposes the conduction channel is depicted with an outer ring of hydrophobic material which is somewhat thicker than the normal width of the hydrocarbon region

of the bilayer. Israelachvili *et al.*^{6,7} and Israelachvili (personal communication) have shown from packing geometry and free energy considerations that this will of necessity lead to a distortion outwards of the bilayer system around the conduction module as indicated in Fig. 2*a*. As pointed out by Israelachvili, the extension of the hydrocarbon chains of the lipids adjacent to the protein module restricts their motion and hence reduces their configurational entropy which leads to surface stress as indicated by the arrows in the diagram.

We now propose that this stress would tend to strain the conduction module radially outwards, which would assist in keeping the sodium channel through this module open. When benzyl alcohol is present we have found that the bilayer hydrocarbon region thickness is increased by at least 25% (somewhat more if the benzyl alcohol increased the dielectric constant of this region). In nerve cells this would have the effect shown in Fig. 2*b*. The hydrocarbon region of the bilayer is now wider than the hydrophobic region of the conduction module, and the bilayer has become distorted inwards around the sodium channel. The surface forces now lead to a radial compression of the channel (see Fig. 2*b*) thus blocking conduction by the channel.

On the other hand, if the conduction module itself stretches to accommodate the changes in the hydrocarbon region of the bilayer, the resulting longitudinal strain could by itself lead to a radial closure, or otherwise interfere with the gating mechanism by way of a deleterious change in the conformation of the protein of the module. Alternatively cardinal sites on it would become buried (if only in the polar head region of the bilayer) as the hydrocarbon region expands. Both these possible effects could readily interfere with the function of many of the proposed excitation mechanisms (for example refs 8, 9).

We thank Dr J. N. Israelachvili of the Australian National University for a stimulating discussion of the packing of proteins in lipid bilayers and for a preprint of his publication on this subject. After submission of this paper we learnt of a manuscript submitted by D. A. Haydon *et al.* in which, on the basis of the effects of alkanes, they postulated a mechanism similar to that we propose here for the molecular basis of anaesthesia.

R. G. ASHCROFT
H. G. L. COSTER
J. R. SMITH

Biophysics Laboratory, School of Physics,
The University of New South Wales,
Kensington, Sydney, Australia 2033

Received 15 March; accepted 22 June 1977.

- 1 Lee, A. G. *Nature* **262**, 545–548 (1976).
- 2 Coster, H. G. L. & Smith, J. R. *Biochim. biophys. Acta* **373**, 151–164 (1974).
- 3 Bell, D. J., Coster, H. G. L. & Smith, J. R. *J. Phys. E*, **8**, 66–70 (1975).
- 4 Ashcroft, R. G., Coster, H. G. L. & Smith, J. R. *Biochim. biophys. Acta* **(469)**, 13–22 (1977).
- 5 Metcalfe, J. C. in *Permeability and Function of Biological Membranes* (eds Bolis, L., Katchalsky, A., Keynes, R. P., Loewenstein, W. B. & Pethica, B. A.) (North-Holland, Amsterdam, 1970).
- 6 Israelachvili, J. N., Mitchell, D. J. & Ninham, B. W. *J. Chem. Soc., Far. Trans. II* **72**, 1525–1568 (1976).
- 7 Israelachvili, J. N. *Biochim. biophys. Acta* **469**, 221–225 (1977).
- 8 Hille, B. *J. gen. Physiol.* **66**, 535–560 (1975).
- 9 Coster, H. G. L. & Smith, J. R. *Bioenergetics Bioelectrochem.* **3**, 214–215 (1976).

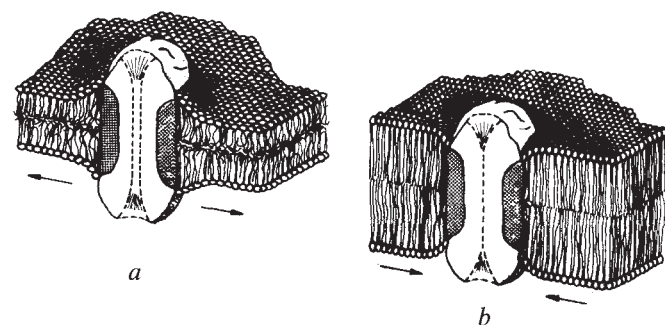


Fig. 2 Schematic representation of the proposed effect of benzyl alcohol on the structure of the bilayer membrane near a hypothetical sodium conduction channel embedded in it. For illustrative purposes, the latter is taken to have a hydrophobic portion (shown shaded) which is somewhat larger than the thickness of the hydrocarbon region of the unperturbed bilayer. The distortion of the bilayer to accommodate such a module together with the requirement of close packing of the fluid-like hydrocarbon portion necessitates the straightening of the fatty acid chains and a closer packing of the polar heads near the conduction module and leads to a radially outward tension on the module. When the hydrocarbon layer thickens in the presence of the anaesthetic, the bilayer is dimpled inwards around the module; this sets up a radial compressive stress on the module. It is suggested that such changes in surface stress are associated with the loss of excitation by the membrane. Alternatively, the accompanying changes in the longitudinal stress in the conduction channel could also lead to inactivation of the gating mechanism.

Quantum amplitude distributions point to functional unity of the synaptic 'active zone'

THE exact nature of the quantum is still not known. Several observations^{1–6} support the idea that transmitter is released from synaptic vesicles. Analyses based on the quantum hypothesis using Poisson^{7,8} or binomial statistics^{9,10} cannot

provide direct evidence for or against this idea, since quanta might equally well be produced by the opening and closing of membrane shutters allowing release of cytoplasmic transmitter. All these considerations rest on the assumption that a single quantum is due to the discharge of a single vesicle or a single shutter action. Recent reports have pointed to a remarkable heterogeneity of the quantum in the mouse and frog neuromuscular junction^{11,12} and in some cases subunits are apparent. In this investigation, evidence is provided from the frog neuromuscular junction that the 'quantum' has subunits.

Spontaneous depolarisations of the postsynaptic cell were first observed in the neuromuscular junction of the frog¹³. Since the amplitude histograms of these depolarisations were well described by a single Gaussian curve they were consequently^{14,15} considered to be due to the liberation of the smallest possible amount of transmitter substance and thus named quanta, single units, and so on. In subsequent experiments (for example, refs 7 and 8) it was found that evoked transmitter release occurs in multiples of these unit packages.

In the present investigation conventional intracellular recording techniques were used to monitor spontaneous miniature endplate potentials (m.e.p.ps) from small muscle cells (20–40 μm diameter) of frog cutaneous pectoris muscles (*Rana temporaria* and *Rana esculenta*) at room temperature.

To rule out systemic errors in measuring the signal amplitudes two independent procedures were used in most experiments. Both involved electronic computing devices (with fast AD converters) which were programmed to recognise the typical signal shape. In the on-line version (PDP8E) a 1-mV signal at the cell corresponded to 231 digital points (thus measuring in steps of $\approx 4 \mu\text{V}$); peak amplitudes were obtained by gliding averaging of the signal trace. In the off-line (PDP10) version, signals were stored on analogue tapes (10 kHz) and later peak amplitudes were evaluated by fitting an idealised theoretical curve to each signal. Both methods produced the same results. Class width and starting point of the histograms were chosen such that the subunits were most clearly visible. Records were carefully investigated for a change in the average amplitude of the m.e.p.ps with time. In the present investigation only stable periods were used for further examinations.

Figure 1 shows amplitude histograms of two different cells. The overall appearance of the distributions is normal. Peaks are, however, apparent at regular distances from one another (dots). To analyse these peaks, a statistical test was adopted which is basically binomial and similar to the methods in use for the classic quantum analysis^{8,10}. It is assumed that there is a constant number of independent subunits (n_v) in each 'active site' (see below), each one having the same probability (p_v) of being discharged when the site is activated. Subunit size (SU = mean amplitude, b = coefficient of variation) is similar in each 'active site' and subunits sum linearly. The size of the subunits as well as the binomial parameters and the predicted curves were estimated by two means.

Method A: Estimation of SU from the amplitude histogram by eye; the multiples have to fit as many as possible of the visible peaks and the first theoretically possible SU must start at 0 mV. The binomial parameters n_v and p_v were then estimated using conventional methods¹⁰. To compare the observed and expected distributions the 'geometrical' method described in detail by Boyd and Martin⁸ was used. b was chosen to give the best fit.

Method B: A non-linear least-squares fit¹⁶, which takes into account all the above assumptions, was used in a computer program to find values for SU , b , n_v and p_v which would provide the best fit to the observed histogram.

Figure 2 shows observed and fitted curves for two

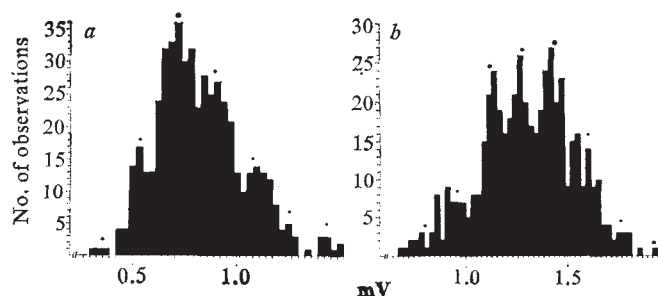


Fig. 1 Intracellular amplitude histograms of spontaneous m.e.p.ps recorded from two different muscle fibres. The total number of observations was 512 in a and 510 in b. Resting membrane potential was 85 mV and 65 mV, respectively. Note peaks at regular distances (dots). The amplitude of the subunits (SU) was estimated using method A. SU is 0.178 mV in a and 0.160 mV in b.

experiments. In all cases there is a good fit thus providing evidence for the existence of subunits. In most experiments (14 out of 18), subunits were apparent, whereas in others no evidence for subunits could be obtained.

Theoretically, the observed subunits could be produced at both the pre- and postsynaptic level or even in the synaptic cleft. Any such mechanism, however, is less plausible than the simple explanation, that spontaneous transmitter release occurs in multi-package units.

In electron micrographs, gatherings of synaptic vesicles are seen at the regularly spaced 'active zones'^{9,17,18}, and it seems that there is a strictly ordered double row with a total of about 10–30 vesicles at each 'active zone'¹⁷. From simultaneous measurements of individual quantal discharges with three micro-electrodes, it has been inferred that spontaneous release of transmitter occurs exclusively at discrete sites ('active sites'), which most likely correspond to the 'active zones'¹². Recent statistical estimations of the number of 'units in the nerve terminal capable of (quantal) release (n)' indicate that in the frog neuromuscular junction, n is small (about 1 per micron length of the nerve terminal). These numbers roughly correspond to the number of 'active zones' in the activated length of the nerve terminal, which in turn suggests that not more than one quantum is discharged at each 'active zone'.

The numbers of subunits in the amplitude histograms were found in several experiments to be between 2 and 15 which are somewhat less than, but still in accord with, the number of vesicles lined up in an 'active zone'. Considering the evidence for a special function of the 'active zone'^{10,12,17}, therefore, we find it plausible that a quantum might be due to the simultaneous discharge of several or all vesicles in one 'active zone'. No temporal resolution of subunits has as yet been observed during a quantal discharge, so a mechanism must be postulated which triggers release from several vesicles virtually simultaneously. In Fig. 2, at some peaks observed frequencies deviate from the expected ones (for example, Fig. 2a, second and third peak of the theoretical distribution). Such deviations merely violate the simple binomial assumption. Indeed, it is more realistic to assume a threshold mechanism such that all vesicles 'available' at an active site are released ($p_v=1$), n_v is changing in time and is different from site to site. The number of subunits observed in the amplitude distributions then shows the range in n_v .

Since the coefficient of variation (b) of the subunits is very small, the releasable transmitter content of individual vesicles must be very similar.

Based on amplitude measurements of a large number of spontaneous m.e.p.ps and application of a statistical test, evidence is presented that quanta consist of about 2 to 15

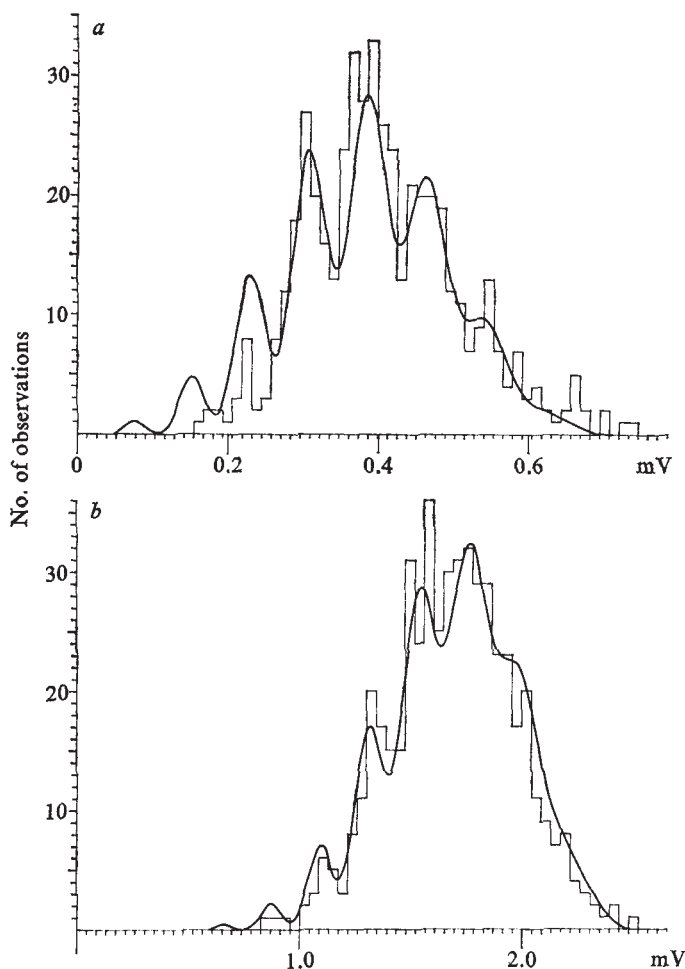


Fig. 2 Observed and fitted amplitude distributions for spontaneous m.e.p.s. from two experiments. Method B was used in *a* and method A in *b*. The amplitude of the subunit (*SU*) was 0.0768 mV in *a* and 0.2177 mV in *b*. The coefficient of variation was 0.11 (*a*) and 0.05 (*b*). The binomial parameter n , was calculated as 8 (*a*) and 10 (*b*). The total number of observations (*N*) was 489 (*a*) and 507 (*b*). The predicted curve starts at 0 mV. Note deviation from the observed distribution especially in *a* (for example, second and third peak of the predicted distribution).

subunits. It is proposed that the quantum is due to the simultaneous discharge of several transmitter quantities at one site. Taking into account other evidence^{10,12,17}, it is postulated that the 'active zone' is a functional unit. Its activation leads to the simultaneous discharge of synaptic vesicles lined up there: this creates a single quantum.

Some of the equipment used in this study was provided for A.W. by the Austrian Fonds zur Förderung der Wissenschaftlichen Forschung.

A. WERNIG
H. STIRNER

Department of Neurophysiology,
Max-Planck-Institut für Psychiatrie,
D-8000 München 40, Kraepelinstrasse 2,
Federal German Republic

Received 18 May; accepted 1 September 1977.

- Palade, G. E. & Palay, S. L., *Anat. Rec.* **118**, 335–336 (1954).
- De Robertis, E. & Bennet, H. S., *J. biophys. Biochem. Cytol.* **1**, 48–58 (1955).
- De Robertis, E., Rodriguez de Lores Arnaiz, G., Salganicoff, L., Pellegrino de Iraldi, A. & Zieler, L. M., *J. Neurochem.* **10**, 225–235 (1963).
- Whittaker, V. P., Michaelson, I. A. & Kirkland, R. J. A., *Biochem. J.* **90**, 293–303 (1964).
- Couteaux, M. R. & Pecot-Dechavassine, M., *C. r. hebdom. Séanc. Acad. Sci., Paris*, **271**, 2346–2349 (1970).
- Weinshilboum, R. M., Thoa, N. B., Johnson, D. G., Kopin, I. J. & Axelrod, J., *Science* **174**, 1349–1351 (1971).
- Del Castillo, J. & Katz, B., *J. Physiol., Lond.* **124**, 560–573 (1954 b).
- Boyd, I. A. & Martin, A. R., *J. Physiol., Lond.* **132**, 74–91 (1956).
- Johnson, E. W. & Wernig, A., *J. Physiol., Lond.* **218**, 751–767 (1971).

- Wernig, A., *J. Physiol., Lond.* **244**, 207–221 (1975).
- Kriebel, M. E., Lladós, F. & Matteson, D. R., *J. Physiol., Lond.* **262**, 553–581 (1976).
- Wernig, A., *Brain Res.* **118**, 63–72 (1976).
- Fatt, P. & Katz, B., *Nature* **166**, 597–598 (1950).
- Fatt, P. & Katz, B., *J. Physiol., Lond.* **117**, 109–128 (1952).
- Del Castillo, J. & Katz, B., *J. Physiol., Lond.* **124**, 553–559 (1954 a).
- Marquardt, D., *J. Soc. Ind. appl. Math.* **11**, 431–441 (1963).
- Couteaux, M. R. & Pecot-Dechavassine, M., *C. r. hebdom. Séanc. Acad. Sci., Paris*, **278**, 291–293 (1974).
- Dreyer, F., Peper, K., Akert, K., Sandri, C. & Moor, H., *Brain Res.* **62**, 373–380 (1973).

Light-activated GTPase in vertebrate photoreceptors

IN studies of the regulation of cyclic GMP levels by light we found that rod outer segment disk membranes contain a cyclic GMP phosphodiesterase (PDE) with a molecular weight of 240,000, whose activation depends on the photoisomerisation of rhodopsin the presence of a nucleoside triphosphate^{1–3}. We report here an additional light-dependent enzyme activity, the hydrolysis of GTP to form GDP and inorganic phosphate, which functions as another enzymatic component in the regulation of photoreceptor guanosine nucleotide levels by light. This GTPase has the action spectrum of rhodopsin and is half-maximally activated by bleaching only one in 2,000 rhodopsins. The enzyme has an apparent K_m for GTP of 0.5 μ M. GTPase activity can be quantitatively removed from the disk membranes and the activity can be fully restored to the disk membranes by a partially purified heat labile protein with an apparent molecular weight of 32,000.

In studies of the activation of photoreceptor cyclic GMP PDE by light and nucleoside triphosphate^{1–3}, we attempted to catalogue all possible reactions of the phosphoryl donors (γ -³²P)GTP and (γ -³²P)ATP. We have observed that between 0.01 and 5 μ M nucleoside triphosphate, light can stimulate the hydrolysis of GTP, but not ATP, in frog (*Rana catesbiana*) outer-segment disk membranes.

The photoreceptor GTPase shows remarkable sensitivity to light; half-maximal activation is observed when only 1 in 2,000 rhodopsins is photoisomerised (Fig. 1). In our conditions (only 0.05% of rhodopsin photoisomerised, 0.5 μ M GTP and 1.2 mg per ml protein), 95% of the radioactivity from (γ -³²P)GTP is found as inorganic phosphate within 10 min and less than 3% covalently labels disk membrane protein (as monitored in sodium dodecylsulphate (SDS)-polyacrylamide gel electrophoresis). The sensitivity of this GTPase to light is identical to that reported for the light-activated photoreceptor PDE (ref. 3).

We next examined the dependence of the activation of GTPase on the wavelength of the exciting light. We arranged conditions where we could compare the efficacy of GTPase activation by equal numbers of photons at wavelengths ranging from 425 to 600 nm. This action spectrum for GTPase corresponds perfectly to the absorption spectrum of the rod photopigment rhodopsin (Fig. 2).

The light activation of photoreceptor GTPase can be accomplished in one of two ways. We can either randomly photoisomerise a small fraction of the total disk membrane rhodopsins (with a flash of light) or we can produce an activation of more than sixfold by adding a 1% admixture of fully photoisomerised membranes to a reservoir of a fully dark-adapted, unwashed suspension of disk membranes. This activation by the phenomenon of mixing was also observed for the light-activated photoreceptor PDE³. This activation is seen in conditions where the contribution of GTPase activity by the 1% admixture is negligible (see Table 1). It is not yet clear whether activation by mixing depends on the fusion of vesicles containing bleached rhodopsin with dark membranes, or whether the bleached material contributes some soluble cofactor in the activation sequence.

The light-activated GTPase activity can be quantitatively removed from the disk membranes by lysing unilluminated rods in 1.0 mM EDTA (pH 7.0), and collecting the disk membranes by centrifugation (35,000 r.p.m., Beckman SW56 rotor, 1 h,

Table 1 Light activation of GTPase by a flash of light or by the combining of dark and fully bleached membranes

	GTPase
0.05% bleached* membranes	300
Dark membranes	48
Dark membranes plus fully bleached membranes† (99:1 protein)	302
Fully bleached component of above mixture	0

GTPase activities are expressed as pmol GTP hydrolysed per mg protein per min. GTP assays were done in the conditions stated in Fig. 1.

*0.05% 'bleached' membranes are disk membrane suspensions in which 0.05% of the rhodopsin is photoisomerised.

†The bleached component of the mixture was washed (by sedimentation and resuspension) three times before full illumination with 100 mM Tris-HCl (pH 7.4) containing 5 mM MgSO₄ and 1 mM dithiothreitol. This washing procedure removes 90% of the GTPase activity.

4 °C). These disk membranes, after one additional wash in the same buffer, showed no GTPase or PDE activities. Both PDE and GTPase activities can be restored to the washed disk membranes by addition of light, 5 mM Mg²⁺ and the supernatant obtained by centrifugation of the lysed rods. This supernatant alone does not show GTPase activity. The supernatant was subsequently applied to a Sepharose 6B column which was equilibrated and eluted with 10 mM Tris (pH 7.5) containing 5 mM MgSO₄, 1 mM dithiothreitol and 50 mM KCl. Column fractions were assayed by recombination with the washed and illuminated disk membranes described above. A single peak of GTPase activity is associated with a heat-labile molecule which does not display PDE activity and which gives a sharp (Coomassie blue positive) band with an apparent molecular weight of 32,000 in SDS-polyacrylamide gel electrophoresis⁶.

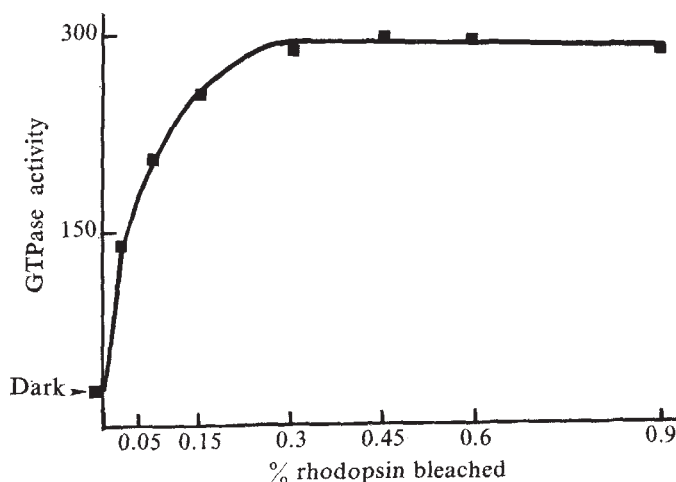


Fig. 1 Light activation of GTPase as a function of the extent of rhodopsin bleaching. Varying degrees of rhodopsin photoisomerisation were obtained by exposing disk membrane suspensions to attenuated (neutral density filtered) light for various times. The filter array was calibrated to permit half-maximal activation with a 30-s exposure. The % rhodopsin photoisomerised was directly measured in 1% hexadecyltrimethylammonium bromide for bleaches in excess of 5% and linearly extrapolated to determine fractional % bleaches accomplished at shorter exposure times. GTPase activity was measured (37 °C at pH 7.4) by the method of Neufeld and Levy⁴ using 0.5 μM (γ-³²P)GTP with specific activity of 20 Ci mmol⁻¹ in a reaction volume of 100 μl containing 90 μg of disk membrane protein. Activities are expressed as pmol GTP hydrolysed per mg protein per min. Alternatively, GTPase was assayed by measuring the formation of GDP from (β,γ-³²P)GTP with descending thin-layer chromatography⁵. Rod outer-segment disk membranes were isolated from *Rana catesbeiana* by sucrose flotation³. All manipulations and GTPase assays (except intentional illumination) were performed in infrared light using image converters³.

We emphasise that the range of GTP concentrations over which light activation is observed is from 0.01 to 5 μM GTP. The K_m for the light-activated reaction is approximately 0.5 μM (±0.4 μM) GTP. The maximum velocity achieved by the light-activated GTPase is approximately 500 pmol GTP hydrolysed per mg protein per min. Above 5 μM GTP, however, no light activation of GTPase is seen but an additional GTPase activity is observed which has a K_m of 90 μM GTP and a V_{max} of 12 nmol GTP hydrolysed per mg protein per min.

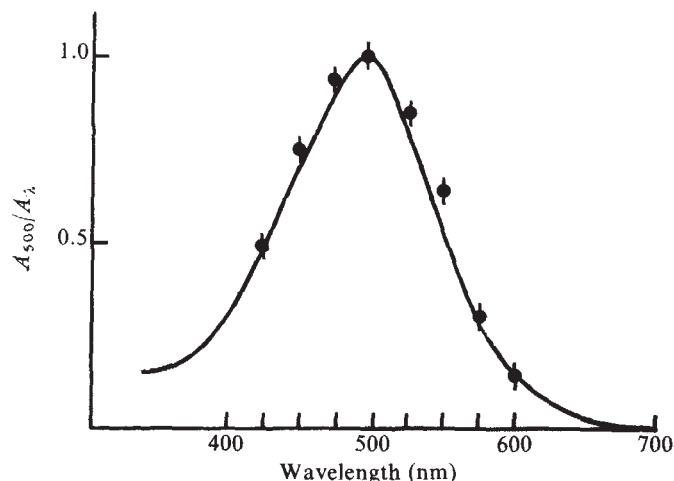


Fig. 2 Action spectrum of light-activated GTPase. Points represent the fraction of highest (A_{500}) GTPase activity above basal. The highest activity observed (at 500 nm) was 80% of maximum possible activation by light. (Basal activity was 30 pmol GTP hydrolysed per mg protein per min). The solid curve is the absorption spectrum of rhodopsin in 1% hexadecyltrimethylammonium bromide. Exposure of disk membranes to equal photon numbers of monochromatic light was accomplished using a light-shielded Gilford 2000 spectrophotometer. The sample holder and photomultiplier were replaced with a camera shutter positioned after the exit slits to allow timed flashes of light. Frosted glass placed in front of the shutter aperture provided a uniformly diffuse source of light. A 100-μl membrane sample was contained in a self-masking cuvette with a 5-mm path length, which was situated after the frosted glass. The intensity of the tungsten lamp at each wavelength was first measured with a YSI radiometer. The intensity profile was used to compute a sample exposure time to normalise photon number at each wavelength. GTPase activities were assayed using the conditions described in Fig. 1.

It is not yet clear whether the high- K_m GTPase activity resides in a second enzyme or reflects changes in the low- K_m enzyme which are brought about by higher (possibly non-physiological) levels of GTP. Interpretation of the physiological roles of the light-activated (low K_m) and non-light-activated (high K_m) GTPase will depend on further studies of GTP concentration *in vivo*.

Finally, we are intrigued by the extraordinary similarity in the light requirements shown by both the GTPase and PDE of vertebrate outer-segment disk membranes. This similarity embraces activation both by a flash of light and by the phenomenon of mixing, suggesting that these enzymes may be functionally, if not topologically, linked. After submission of this manuscript we were informed of a report of similar findings published in the abstracts of the 21st Biophysical Society meeting⁷.

G. L. WHEELER
Y. MATUO
M. W. BITENSKY

Yale University School of Medicine,
Department of Pathology,
New Haven, Connecticut 06510

Received 18 July; accepted 29 August 1977.

1. Bitensky, M. W. *et al. Adv. Cyclic Nucleotide Res.* **5**, 213–240 (1975).
2. Miki, N., Baraban, J. M., Keirns, J. J., Boyce, J. J. & Bitensky, M. W. *J. biol. Chem.* **250**, 6320–6327 (1975).
3. Keirns, J. J., Miki, N., Bitensky, M. W. & Keirns, M. *Biochemistry* **14**, 2760–2766 (1975).
4. Neufeld, A. & Levy, H. *J. biol. Chem.* **244**, 6493–6497 (1969).
5. Keirns, J. J., Wheeler, M. A. & Bitensky, M. W. *Analyt. Biochem.* **61**, 336–348 (1974).
6. Fairbanks, G., Steck, T. L. & Wallach, D. F. H. *Biochemistry* **10**, 2606–2617 (1971).
7. Robinson, W. E. & Hagins, W. A. *Biophys. J.* **17**, 196A (1977).

Salt-sensitive *in vitro* protein synthesis by a moderately halophilic bacterium

EXTREMELY halophilic bacteria, such as *Halobacterium cutirubrum*, grow only in high NaCl concentrations (2.5–5 M) and have very high (5 M or higher) internal concentrations of salts, mainly KCl^{1,2}. The ribosomes and protein-synthesising systems of these organisms seem especially adapted to function in such concentrations³. In contrast, moderately halophilic bacteria, such as *Vibrio costicola*, grow over a much wider range of NaCl concentrations (at least 0.5–3.5 M)⁴. In *V. costicola*, though not necessarily in all moderate halophiles, the cell-associated monovalent cations are at least as concentrated as those of the external medium. For example, cells growing exponentially in medium containing 1 M NaCl and 0.008 M KCl can have internal Na⁺, K⁺, and NH₄⁺ concentrations of about 0.6, 0.7 and 0.4 M respectively, as well as 40 mM Mg²⁺ (ref. 5 and our unpublished results). Ribosomes from *V. costicola* differ from those of both extremely halophilic and non-halophilic bacteria in their ability to maintain a 'standard' sedimentation pattern (30, 50 and 70S) over a wide range of salt concentrations. This pattern is not changed by the NaCl concentration in which the cells are grown⁶. Such properties, however, do not really tell us how well the ribosomes function at different salt concentrations. Studies of *in vitro* protein synthesis, reported here, suggest that ribosomes may function at much lower salt concentrations than measurements of total cell-associated ions indicated to be present in the cell.

Extracts were prepared from cells grown in medium containing 1 M NaCl, washed, and suspended in mixtures of NaCl, KCl, NH₄Cl and MgCl₂ approximating the cell-associated concentrations of these cations. Cells were broken by osmotic lysis or by grinding with alumina. After centrifugation at 30,000g for 30 min to remove intact cells and membrane fragments, the supernatants (S-30) were incubated with the ingredients used to demonstrate *in vitro* protein synthesis in other bacterial extracts (see legends to Figs 1 and 2), but with the final salt concentrations adjusted to that in which the cells were broken. No activity was found, as measured by incorporation of radioactive amino acids into trichloroacetic acid-precipitable material. But, if no salt was added to the assay mixture other than that contained in the extract (final concentration about 0.2 M) protein synthesis occurred. If the concentration of any salt was increased, activity fell. At concentrations of 1 M or higher, no activity could be measured.

Figure 1 shows the effects of various concentrations of salts of monovalent cations and MgCl₂ on poly U-dependent phenylalanine incorporation. Figure 2 gives comparable results for incorporation using 'endogenous messengers'—that is, incorporation of a mixture of ¹⁴C-labelled amino acids from a *Chlorella* hydrolysate. In both systems activity was highest in NH₄Cl, followed by KCl, then NaCl. Optimal activities occurred in 0.1–0.2 M salts, and higher concentrations were inhibitory. This was especially noticeable in the effects of NaCl on the 'endogenous' system. Magnesium ions were needed for activity, maximal activity being reached at about 18 mM Mg²⁺ in both systems.

We considered the possibility that high concentrations of monovalent cations inhibited protein synthesis because they were displacing the essential divalent cation, Mg²⁺. If this were so, it should be possible to overcome the inhibitory effect of the former by increasing the latter. In both systems, however, increasing the MgCl₂ concentration two- or fivefold (from 20 mM) did not reverse the inhibition caused by higher concentrations of the other salts (up to 1 M). Higher concentrations of Mg²⁺ were just as inhibitory in the presence of higher as in lower concentrations of the other salts.

One of the problems with establishing a protein synthesising system in a 'new' bacterium, is showing that the measured incorporation represents polypeptide synthesis. This was shown by the following findings. Incorporation is energy dependent. In the absence of poly U phenylalanine incorporation was reduced by 90%, and ¹⁴C-leucine was incorporated to about the same extent as phenylalanine. Leucine incorporation was not stimulated by poly U. Treatment of the system with ribonuclease A (Sigma, 40 µg ml⁻¹ in assay mixture) abolished all activity. Chloramphenicol (15 µg ml⁻¹) and puromycin (25 µg ml⁻¹) strongly or completely inhibited activity. Aurintricarboxylic acid (105 µg ml⁻¹) inhibited incorporation in both the poly U-dependent

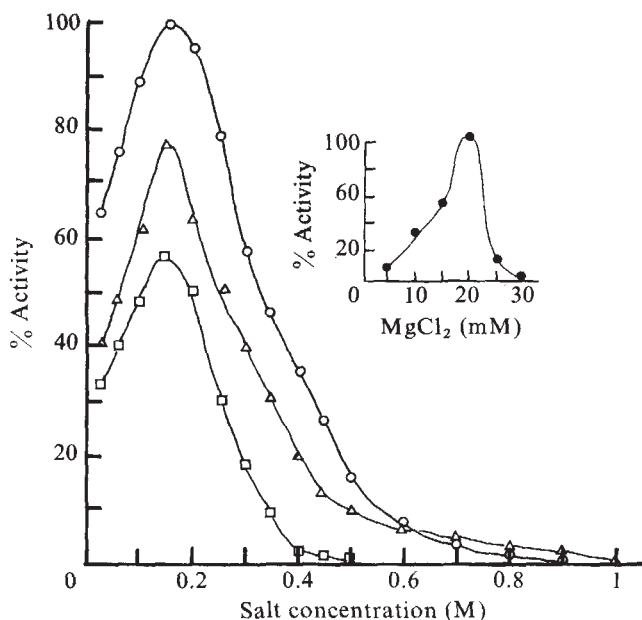


Fig. 1 Effects of NaCl (□), KCl (△), NH₄Cl (○) and MgCl₂ (insert) on the poly U-dependent ¹⁴C-phenylalanine incorporation. *V. costicola* cells, grown in 1 M NaCl, were broken by grinding with twice the wet cell weight with alumina. After centrifugation at 30,000g for 30 min to remove whole cells and membrane fragments, small volumes (1–2 ml) of the extracts were dialysed with rapid stirring for 1 h against 1–2 l of a solution of 100 mM NaCl, KCl, or NH₄Cl, depending on the salt to be studied, plus 20 mM MgCl₂ and 10 mM Tris-HCl, pH 7.6. This dialysed extract was then centrifuged at 30,000g for 30 min and the resulting supernatant was termed the S-30 extract. 90 µl of the S-30 extract was incubated with 210 µl of the reaction mixture so that the final concentrations were: 15 mM phosphoenol pyruvate; 2 mM adenosine triphosphate; 1 mM guanosine triphosphate; 18 mM MgCl₂ (or as indicated, insert); 30 µg ml⁻¹ pyruvate kinase; 50 mM Tris-HCl, pH 7.6; 7.5 mM reduced glutathione; 9.6 µM ¹⁴C-phenylalanine (522 mCi mmol⁻¹); 270 µg poly U; and 100 mM NH₄Cl, KCl, or NaCl, or the concentration as indicated. The reaction was incubated at 30 °C and 50 µl aliquots taken at 0, 2, 4, 6, and 8 min. The aliquots were precipitated with 3 ml of cold 5% (w/v) trichloroacetic acid, heated at 90 °C for 20 min, then collected onto 0.45-µm filters (Millipore). The filters were dried and counted in a liquid scintillation counter with 5 ml of Scintiverse (Fisher Scientific). An activity of 100% represents 22.2 pmol of phenylalanine incorporated per min per A₂₆₀ unit of the S-30 extract.

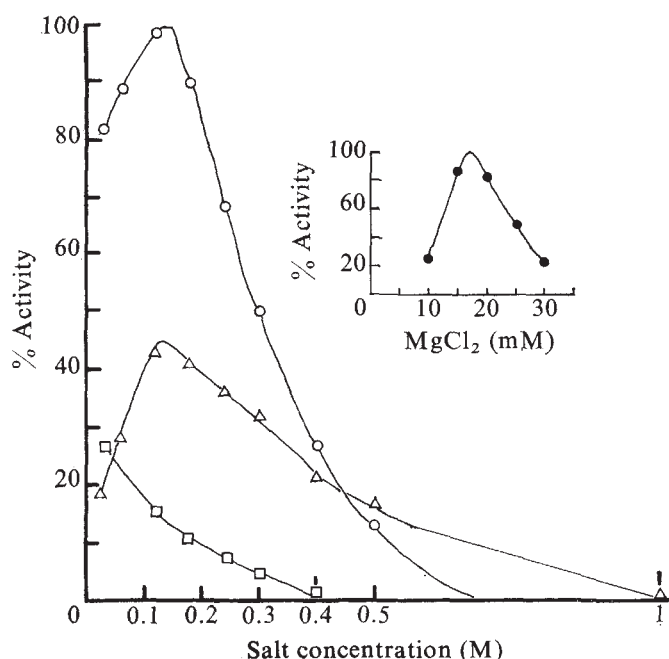


Fig. 2 Effects of NaCl (○), KCl (△), NH₄Cl (□) and MgCl₂ (insert) on the amino acid incorporation dependent on 'endogenous messenger'. *V. costicola*, grown in 1 M NaCl, cells were lysed osmotically by suspending the cell pellets in buffer containing: 200 mM KCl; 10 mM Tris, pH 7.6; and 5 mg ml⁻¹ lysozyme. After 2–3 min incubation on ice, the ionic concentration and composition was restored to: 100 mM KCl; 10 mM Tris, pH 7.6; and 18 mM MgCl₂. S-30 extracts were prepared and dialysed as in Fig. 1. The assay conditions are as in Fig. 1 except that poly U and ¹⁴C-phenylalanine were omitted and ¹⁴C-labelled *Chlorella* hydrolysate (54 mCi per mAtom of carbon) was substituted. The values are expressed as % of maximum activity, maximum activity being 300 c.p.m. per min per A₂₆₀ unit.

and the endogenous system. Analysis of the end products of ¹⁴C-phenylalanine incorporation by the method of Pestka *et al.*⁷ showed that 58% was in the form of diphenylalanine, 25% in the form of triphenylalanine, and 17% in the form of higher oligo- or polypeptides.

Salts had a different effect on the stability of the protein synthesising system than on its activity (Table 1). Concentrations of all monovalent salts which inhibited protein synthesis could stabilise the system. Usually, the system was more stable at lower temperatures. Only in the highest NaCl concentration (2 M) did rapid inactivation of the system occur (not shown in Table 1). We cannot assign the effects of the salts to any specific component of the protein synthesising system, but it has been shown that certain salts

Table 1 Effects of Na⁺, K⁺, and NH₄⁺ chlorides on the stability of the poly U-dependent incorporation of ¹⁴C-phenylalanine

Salt	<i>t</i> _{1/2} values (min)		
	30 °C	20 °C	3 °C
0.1 M NH ₄ Cl	22	160	510
1.0 M NH ₄ Cl	63	∞	∞
0.1 M KCl	40	57	405
1.0 M KCl	20	∞	∞
0.1 M NaCl	30	45	∞
1.0 M NaCl	48	58	1,095

400 μl of the S-30 extract, dialysed against 100 mM of the appropriate salt was incubated with solid salt to raise the concentration to the desired level. Aliquots of 90 μl of this mixture were removed at various times and assayed as described in Fig. 1 except that for the 1 M salt concentrations no monovalent ions were added to the reaction mixture. This resulted in a final concentration of 0.3 M salt. The *t*_{1/2} is the time required for the extract to lose half of the original activity. The ∞ symbol indicates that the extract lost no activity over the test period (usually 8 h).

have different effects on the stability and the activity of individual enzymes of halophilic bacteria^{2,8}.

It is very difficult to predict the state of ions in the cytoplasm of a cell from knowledge of the total concentration alone. Certain physiological functions might provide a better estimate of the actual internal state of these ions. It is probably significant that many of the internal enzymes of the halobacteria function better in KCl than in NaCl^{2,8}. It is more significant that protein synthesis, a process dependent on many enzymes and on the proper configurations of ribosomes and their interactions with mRNA, also functions in the conditions that are suggested to exist in these cells by total ionic analysis. This is not so for *V. costicola*. Rather, protein synthesis occurs in much lower salt concentrations than those found in the cell as a whole. Protein synthesis would be completely inhibited by the salt concentrations found in the bacteria, if all were free in the cytoplasm. Our results suggest that they are not free, but rather that they are sequestered within the cell. Possibly Mg²⁺ is bound to the cell envelope (optimum Mg²⁺ level for protein synthesis is less than half of total cellular concentration). Many bacterial envelopes can bind divalent cations⁹. It is less easy to envisage a binding site for the monovalent cations. A binding of such cations to envelopes has been reported for other moderately halophilic and non-halophilic bacteria, however (discussed in ref. 5). We are investigating this possibility for *V. costicola*.

This work was performed when T.H. was a visiting scientist at Ottawa University.

R. M. WYDRO

Department of Biology,
University of Ottawa,

W. MADIRA

Department of Biochemistry,
King's College,

T. HIRAMATSU

Department of Bacteriology,
Osaka City University Medical School,
Osaka, Japan

M. KOGUT

Department of Biochemistry, King's College,
Strand, London WC2, UK

D. J. KUSHNER

Department of Biology, University of Ottawa,
Ottawa, Ontario, K1N 6N5, Canada

Received 17 May; accepted 6 September 1977.

1. Kushner, D. J. *Adv. Appl. Microbiol.* **10**, 73 (1968).
2. Brown, A. D. *Bact. Rev.* **40**, 803 (1976).
3. Bayley, S. T. in *Extreme Environments: Mechanism of Microbial Adaptation* (ed. Heinrich, M. R.) 119 (Academic, New York, 1976).
4. Forsyth, M. P. & Kushner, D. J. *Can. J. Microbiol.* **16**, 253 (1970).
5. Shindler, D. B., Wydro, R. M., & Kushner, D. J. *J. Bact.* **130** (1977).
6. Wydro, R., Kogut, M., & Kushner, D. J. *FEBS Lett.* **60**, 210 (1975).
7. Pestka, S., Scolnick, E. M., & Heck, B. H. *Analyt. Biochem.* **28**, 376 (1969).
8. Lanyi, J. K. *Bact. Rev.* **38**, 272 (1974).
9. Beveridge, T. J. & Murray, R. G. E. *J. Bact.* **127**, 1502 (1976).

Carbon dioxide governs the oxygen affinity of crocodile blood

OXYGEN affinity of haemoglobin inside red blood cells is decreased to a level where oxygen unloading can proceed at sufficiently high partial pressures in the tissue capillaries. This adaptation of oxygen affinity is accomplished mainly by the interaction of haemoglobin with intraerythrocytic phosphates¹. These comprise 2,3-diphosphoglycerate (DPG) in most mammals, ATP and GTP in fish², and *myo*-inositol pentaphosphate (IP₅) in mature birds^{3–5}. No haemoglobin system has been described in higher vertebrates which is devoid of functional interaction with the respective intracellular phosphates. Recently, however, Sullivan⁶ has questioned the role of phosphates for the control of oxygen affinity in crocodile

red cells (ATP and inorganic phosphate have been identified^{3,7,8}). He remarks that "the substance regulating oxygen affinity in alligators is very efficient, but unknown". We show here that the oxygen affinity of haemoglobin from a member of the order crocodiles (*Crocodylus porosus*) is not affected by any of the regulating allosteric cofactors known with the exception of molecular CO₂ and protons.

A specimen of *C. porosus*, about 5-yr-old, was anaesthetised with Rompun and Katalar and blood obtained by heart puncture. Red cells were washed in saline and lysed by addition of distilled water. The ghosts were removed by centrifugation. The clear supernatant was passed over Sephadex G-25 columns at 4 °C equilibrated with Tris buffer pH 8.1 to remove all phosphates, and subsequently dialysed against 0.1 M NaCl. The oxygen affinity of haemoglobin solutions was determined as described earlier⁹.

The partial pressure of oxygen at which purified haemoglobin solution is half saturated with oxygen (P_{50}) was 3.9 Torr (20 °C, pH 7.3) and increased to 28.8 Torr on addition of CO₂ ($p\text{CO}_2 = 40$ Torr, pH 7.3). This value is very near to the one found in whole blood ($P_{50} = 30.3$ Torr at 20 °C, extracellular pH 7.47, $p\text{CO}_2 = 40$ Torr). At more alkaline pH the effect of CO₂ on P_{50} is even more pronounced (Fig. 1). Note that the concentration of ATP was low in crocodile erythrocytes (0.6 mM) and that UTP, CTP, ITP, GTP and DPG could not be detected enzymatically¹⁰ or by thin-layer chromatography². Apart from ATP the following phosphates were

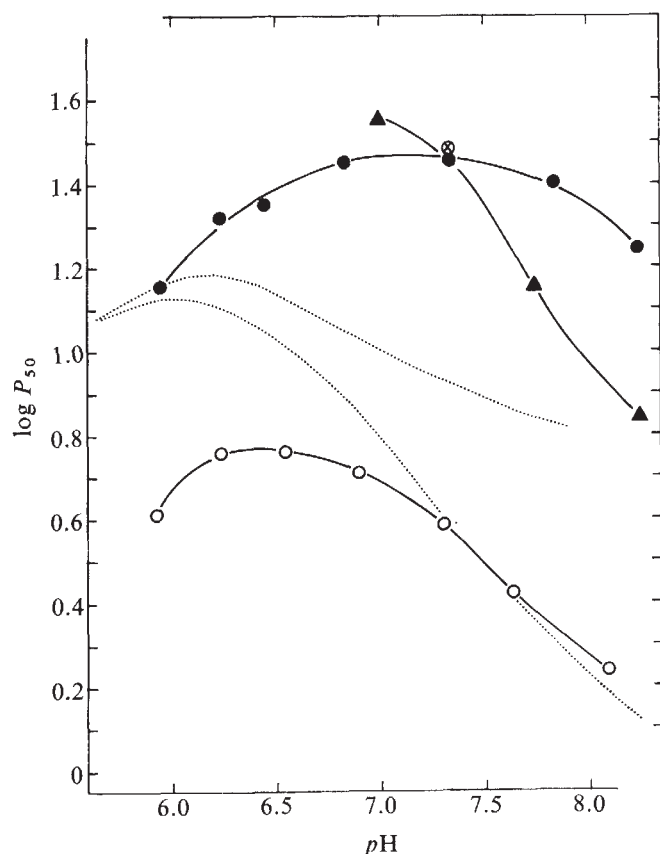


Fig. 1 Plot of $\log P_{50}$ against pH of crocodile haemoglobin in the absence (○) of CO₂ and at a $p\text{CO}_2$ of 40 Torr (●) at 20 °C. Hb concentration 120 μM haem; buffers: 50 mM bis-Tris-100 mM NaCl in the absence of CO₂ (total ionic strength 0.150) and bicarbonate-NaCl buffer in the presence of CO₂, 150 mEq l⁻¹. pH was also varied by changing $p\text{CO}_2$ at constant bicarbonate concentration (▲). Blood P_{50} was determined by means of the Lex-O₂-Con apparatus at 20 °C, $p\text{CO}_2$ 40 Torr and extracellular pH 7.47 (⊗). The dotted lines were obtained from similar experiments on human haemoglobin in the absence (lower) and presence (upper) of CO₂.

tested with respect to their possible influence on P_{50} of haemoglobin solutions: IP₅ prepared from chicken blood¹¹, DPG, fructose-1,6-diphosphate, and cyclic AMP, each used in a 10–20-fold excess of phosphate over Hb₄. None of these led to a significant shift of P_{50} . Similarly, inorganic phosphate (0.1 M) had no effect on the oxygen affinity of crocodile haemoglobin. Apart from CO₂ only protons led to heterotropic interactions with the oxygen binding: $\Delta \log P_{50} / \Delta \text{pH} = -0.43$ in the alkaline range (Fig. 1). When the change in pH was produced by varying $p\text{CO}_2$ at constant HCO₃⁻ concentration $\Delta \log P_{50} / \Delta \text{pH}$ increased to -0.66 due to the additional action of CO₂ on P_{50} . Cooperativity of oxygen binding was not significantly influenced by CO₂: the exponent n in Hill's equation¹² was 2.1 in the absence and 2.0 in the presence of CO₂.

The question arises then as to how this remarkable effect of CO₂ on the oxygen affinity of crocodile haemoglobin can be explained on a molecular basis. As reptile haemoglobins polymerise in certain conditions^{6,13} it is conceivable that the action of CO₂ leads to increased aggregation with consecutive decrease in oxygen affinity. Gel filtration experiments on Sephadex G100 columns operated with bis-Tris buffer pH 7.3 at 20 °C yielded two components in equal proportions, one which is probably 'octameric' (apparent molecular weight 110,000) and one 'tetrameric' species (apparent molecular weight 43,000). This proportion was very much the same when the experiment was carried out with bicarbonate buffer equilibrated continuously with CO₂ so to give the same pH and ionic strength. Furthermore, oxygen binding curves of the isolated components showed that CO₂ produced the same decrease in oxygen affinity in the 'octameric' and in the 'tetrameric' pigment as in the unfractionated haemolysate. Additional control experiments on Sephadex G100 were carried out to check that the respective components had not interconverted during the oxygen binding experiments. Both the 'octameric' and the 'tetrameric' component appeared as one major band in isoelectric focusing experiments with isoionic pH values of 7.24 and 7.26, respectively.

Thus we conclude that a reversible aggregation-dissociation phenomenon is not the cause for the CO₂-induced decrease in oxygen affinity of crocodile haemoglobin. Most likely the formation of oxygen-linked carbamate compounds is responsible for the observed effects, qualitatively similar to what is found in human haemoglobin, where more carbamate is present at the terminal α-amino groups in the deoxy- than in the oxy-structure. If the same residues mediate the carbamate formation in crocodile haemoglobin, our data suggest that the carbamate equilibrium constant as well as the change of the carbamate equilibrium constant when the haemoglobin binds or releases oxygen must be higher than even in the β-subunits of human haemoglobin, where most oxygen-linked carbamate is being formed^{9,15}. The negatively-charged carbamate itself could stabilise the deoxy-structure and de-stabilise the oxy-structure of crocodile haemoglobin by secondary interactions with positively-charged residues. If such interactions lead to structural transitions of the quaternary structure it may be possible to observe changes in the absorption spectrum as, for example, in certain fish haemoglobins at different pH values¹⁶. Difference spectra in the Soret range obtained on crocodile haemoglobin in the absence and presence of CO₂ ($p\text{CO}_2 = 100$ Torr) gave no indication of such transitions, however.

Whatever the exact molecular mechanism of the strong antagonism between CO₂ and oxygen binding turns out to be, it is significant that this particular interaction, which must have appeared quite early in the evolution of haemoglobin function in vertebrates^{14,17}, has been explored very efficiently in a member of the order *Crocodylia* which possesses haemoglobins with a high intrinsic oxygen affinity and only little intra-erythrocytic phosphates. In situations where the animal moves about very quickly and therefore liberates CO₂ from the tissues, the antagonism between CO₂ and oxygen binding should be very efficient in aiding oxygen release from the blood.

We thank Dr Wiesner from the Tierpark Hellabrunn,

München, for blood samples, Mrs U. Engels for technical assistance and to Dr G. Gros for helpful comments.

Note added in proof: Oxygen binding to haemoglobin from another crocodile species, *Caiman crocodylus*, exhibits the same strong interaction with molecular CO₂, but none with ATP.

CHRISTIAN BAUER
WOLFGANG JELKMANN

Physiologisches Institut der Universität
8400 Regensburg, West Germany

Received 4 July; accepted 22 August 1977.

1. Benesch, R. E. & Benesch, R. *Adv. Prot. Chem.* **28**, 211–237 (1974).
2. Weber, R. E., Lykkeboe, G. & Johansen, K. *J. exp. Biol.* **64**, 75–88 (1976).
3. Rapoport, S. & Guest, G. *J. biol. Chem.* **138**, 269–282 (1941).
4. Johnson, L. F. & Tate, M. E. *Can. J. Biochem.* **47**, 63–73 (1969).
5. Isaacs, R. E. *et al. Comp. Biochem. Physiol.* **55A**, 29–33 (1976).
6. Sullivan, B. in *Chemical Zoology* **9**, (eds Florkin, M. & Scheer, B. T.) 377–398 (Academic, New York, 1974).
7. Bartlett, G. R. *Comp. Biochem. Physiol.* **55A**, 207–210 (1976).
8. Coates, M. J. *mol. Evol.* **6**, 285–307 (1975).
9. Bauer, C., Baumann, R., Engels, U. & Pacyna, B. *J. biol. Chem.* **250**, 2173–2176 (1975).
10. Bergmeyer, H. U. (ed.) *Enzymatische Methoden* **2** (Verlag Chemie, Weinheim, 1974).
11. Bartlett, G. R. & Borgese, T. A. *Comp. Biochem. Physiol.* **55A**, 207–210 (1976).
12. Hill, A. V. J. *Physiol., Lond.* **40**, 4–7 (1910).
13. Sullivan, B. & Riggs, A. *Biochim. biophys. Acta* **140**, 274–283 (1967).
14. Bauer, C. & Kurtz, A. *J. biol. Chem.* **252**, 2952–2955 (1977).
15. Morrow, J. S., Matthew, J. B., Wittebort, R. T. & Gurd, F. R. *J. biol. Chem.* **251**, 477–484 (1976).
16. Giardana, B., Ascoli, F. & Brunori, M. *Nature* **256**, 761–762 (1975).
17. Bauer, C., Engels, U. & Paleus, S. *Nature* **256**, 66–68 (1975).

Hapten-induced allosteric transition in the light chain dimer of an immunoglobulin

THE dimer of immunoglobulin light chains (L₂) was shown by X-ray crystallography to be structurally homologous to the native immunoglobulin Fab fragment (HL) (refs 1–3). This led to the suggestion that L₂ is a model for a primitive antibody¹ and for the T-cell receptor^{22,23}. In L₂ 315 (from the mouse myeloma protein MOPC 215) functional homology was also implicated in that both L₂ and HL bind the same nitroaromatic haptens with similar fine specificity^{4,5}. We have considered the possibility that this homology extends also to hapten-induced conformational changes and that studies with L₂ will yield information on the nature of such processes in immunoglobulins. L₂ 315 constitutes a particularly good model system for such studies as it binds two hapten molecules per dimer at symmetry-related sites^{4,5}. It is thus possible to monitor the conformational effects of the first binding hapten through the mode of interaction with the second, as revealed by the shape of the saturation curve⁶. We report here that dinitrophenyl-lysine (DNPL) and its fluorescent analogue nitrobenzoxadiazole-alanine (NBDA) (ref. 7) both display sigmoidal binding curves, implying positive cooperativity in their interaction with L₂ 315. This observed cooperativity is shown to arise from a hapten-induced conformational transition which may be described by the allosteric model⁶. We propose that this allosteric transition involves changes in the relative position of the chains in L₂, lending support to suggestions^{8,9} that a similar process occurs in the structurally homologous intact immunoglobulin. These findings are, therefore, relevant to the question of antigen-triggered effector functions of antibodies through an allosteric mechanism^{8,10–12}.

The difference spectra observed on addition of hapten to L₂ 315 are shown in Fig. 1a and b. Each hapten displays its characteristic difference spectrum which varies little throughout the titration as is particularly emphasised in the well defined isosbestic points for NBDA. This strongly suggests that the two sites on L₂ 315 are identical with respect to hapten binding. Accordingly, any model used to account for the observed cooperativity should be symmetric, in line with the previously suggested local symmetry around the binding site of L₂ 315 (ref. 5).

The hapten occupancies (calculated from the absorption changes) were plotted against hapten concentration (Fig. 2). For both haptens the binding curve is sigmoidal, the effect being more pronounced with NBDA. For this hapten equi-

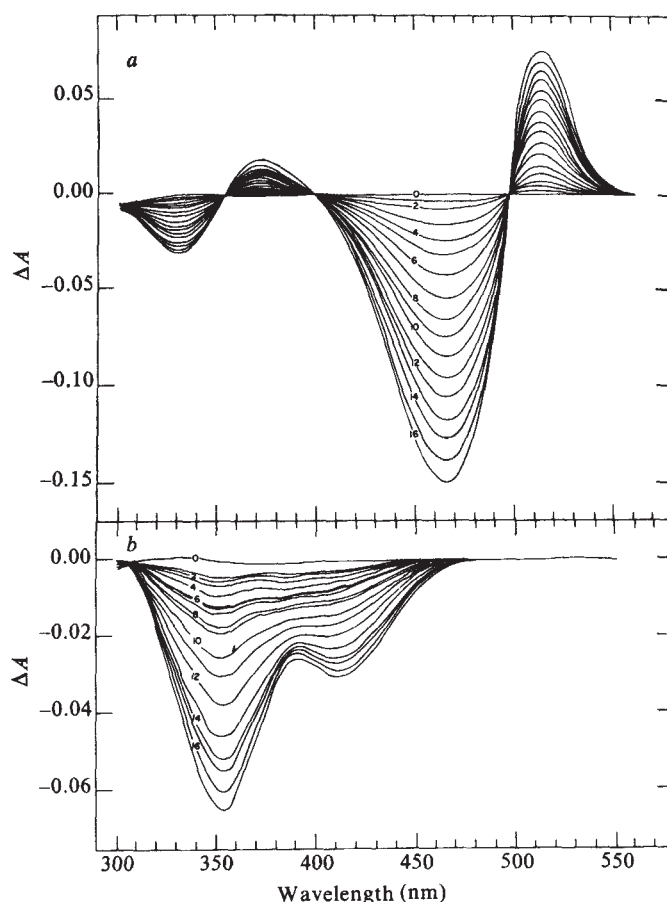


Fig. 1 Difference spectroscopic titrations in which *a*, 4-(α -N-L-alanine)-7-nitro-benz-2-oxa-1,3-diazole (NBDA)⁷, and *b*, ϵ -N-2,4-dinitrophenyl-L-lysine (DNPL), are added to mildly reduced and alkylated L₂ 315⁴ in 0.01 M sodium phosphate buffer pH 7.4 with 0.15 M NaCl (PBS). ΔA is the change in hapten absorption on binding to L₂ ($\Delta A = A_{\text{bound}} - A_{\text{free}}$). Spectra were taken on a double beam Cary 118 spectrophotometer in rectangular quartz cuvettes of 0.437 cm pathlength at 4 °C. Equal aliquots of hapten solution were added to a cuvette containing the protein solution and to a reference cuvette containing an exactly equal volume of buffer. The minor contribution of protein absorption below 320 nm was corrected electronically. Initial L₂ concentration was 6.0×10^{-5} M sites (assuming 2 sites per protein molecule, and $\epsilon_{280} = 60,000 \text{ M}^{-1} \text{ cm}^{-1}$ for the dimer).

brum dialysis measurements were also performed and the points were found to lie on the same binding curve (Fig. 2). The origin of the observed sigmoidity is therefore thermodynamic and not spectroscopic—that positive cooperativity occurs in the binding of hapten. The data were analysed according to the Adair¹³ formalism as described in Table 1.

The underlying mechanisms of the observed positive cooperativity may be one of the following: (1) the first bound hapten exerts a direct attractive interaction on the second; (2) hapten binding is coupled to a chain-dissociation equilibrium of the form $L_2 \rightleftharpoons 2L$ (refs 14, 15); (3) hapten binding is linked to an allosteric conformational transition in L₂. Since the sites on L₂ 315 are probably no more than 12 Å apart⁵, direct contact interaction between haptens is possible. But there are strong arguments against it, which are: first, from the chemical nature of the ligands no mutual attraction is expected, particularly not for the negatively charged NBDA; second, attraction between identical ligands would be fortuitous and is unlikely to occur in two different cases; and third, direct hapten interaction is expected to lead to spectral asymmetry because the first hapten interacts only with protein residues while the second hapten also interacts with the first. The possibility of chain dissociation at the L₂ concentration studied (3×10^{-5} M) was ruled out by a series of sedimentation

Table 1 The best fit parameters for the curves in Fig. 2 according to different models

Hapten	Intrinsic Adair constants (M^{-1})	n_H	λ (nm)	$\Delta\epsilon_\lambda$ ($cm^{-1}M^{-1}$)	L	Intrinsic MWC constants (M^{-1})	$c = K_T/K_R$
NBDA	$K_1 = 5.0 \times 10^2$ $K_2 = 5.6 \times 10^4$	1.83	464	-12400	110	$K_T = 0$ $K_R = 5.6 \times 10^4$	0
DNPL	$K_1 = 3.5 \times 10^3$ $K_2 = 1.4 \times 10^4$	1.33	354	+6800 -4500	110	$K_T = 2.9 \times 10^3$ $K_R = 6.7 \times 10^4$	0.043

The Adair formalism¹³ makes no mechanistic assumptions other than that L_2 has two initially identical sites, and it enables us to calculate their apparent consecutive binding constants, K_1 , K_2 (see ref. 14). The binding curves are fitted according to this formalism assuming a constant value of $\Delta\epsilon$ throughout the titration (see text). λ is the wavelength at which $\Delta\epsilon$ is measured. n_H is the slope of the Hill plot (insert of Fig. 2) at half saturation, where $\log [R/(N-R)] = 0$. It may be calculated as $n_H = 2/[1 + (K_1/K_2)^{1/2}]$, and for positive cooperativity $1 < n_H \leq 2$ (ref. 14). The MWC model⁶ assumes an equilibrium between two conformations of the hapten-free protein $T_0 \rightleftharpoons R_0$ ($L = T_0/R_0$) which bind the hapten with different association constants K_T , K_R ($c = K_T/K_R < 1$). The MWC parameters were calculated from the Adair constants according to ref. 14. Both exclusive binding ($K_T = 0$) and non-exclusive binding ($K_T > 0$) are mathematically equivalent to the Adair scheme and are indistinguishable by parameter fit. Exclusive binding is the simplest as it makes possible calculation of unique values for L and K_R . As L is an attribute of the protein, it should be equal for both haptens. It was found impossible to fulfill this unless non-exclusive binding is assumed either for DNPL or for both haptens. We chose the first possibility thus getting a lower-limit estimate of L . Hapten binding has an appreciable effect on the conformational equilibrium: in the absence of hapten only about 1% is in the R form [$1/(1+L)$] while in the fully saturated protein R amounts to 100% and 83% for NBDA and DNPL respectively [$1/(1+Lc^2)$]. This is in line with the fact that the free energy change in the T to R transition is $\Delta G = -RT \ln(1/L) = 2.8 \text{ kcal mol}^{-1}$, which is comparable to $-RT \ln K_R \sim -6.5 \text{ kcal mol}^{-1}$ for the binding of both DNPL and NBDA.

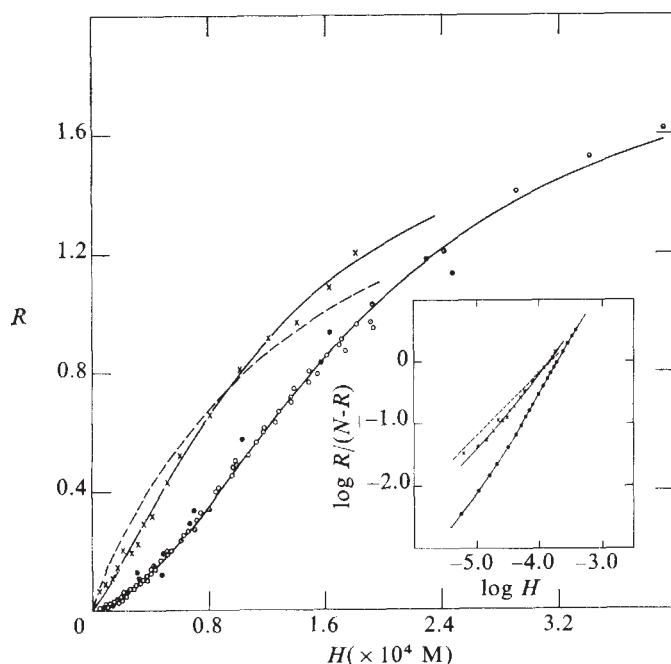


Fig. 2 Saturation curves of L_2 315 with NBDA: $\circ$, ΔA at 464 nm; $\odot$, ΔA at 515 nm and with DNPL; $\times$, ΔA at 354 nm. R is the average site occupancy by the hapten, calculated as $R = \Delta A/(\Delta\epsilon A_0)$ where A_0 is the molar concentration of L_2 and $\Delta\epsilon = \epsilon_{bound} - \epsilon_{free}$ is the change in the molar extinction of the hapten on binding ($R < 2$). H is free hapten concentration. For NBDA equilibrium dialysis points ($\bullet$) are also shown. These were measured as described⁵ and concentrations were monitored spectrophotometrically. A parameter fitting nonlinear least square procedure²¹ was used to obtain the values of K_1 , K_2 and $\Delta\epsilon$, as given in Table 1. The solid lines are drawn using these values of the parameters. The broken line is calculated assuming a single association constant $K = 6.3 \times 10^3 M^{-1}$ as previously reported⁴ for the binding of DNPL to L_2 315. It is seen that the difference between the solid and broken lines for DNPL is relatively small which may explain why in a previous (equilibrium dialysis) study with this hapten the cooperativity remained undetected. It should be stressed that in view of the observed cooperativity any previous measurements in which the binding curve was not extensively characterised should be treated with caution. Insert: the saturation points replotted in the form of a Hill plot. N , the number of binding sites per L_2 is 2.0. For NBDA only one titration (ΔA at 515 nm) is reproduced for clarity. The lines are drawn with the same calculated parameters as in the main figure. The slope at $\log [R/(N-R)] = 0$ is denoted n_H and its value is given in Table 1.

measurements in which a value of $S_{20,w} = 3.3 \pm 0.15$ (corresponding to L_2) was found in the concentration range of $6.0 \times 10^{-6} M$ to $2.0 \times 10^{-4} M$ in the presence or in the absence of hapten.

A conformational change induced in L_2 on hapten binding seems therefore to be the most plausible explanation for the observed positive cooperativity. The constancy of spectral effects through the titration strongly supports the notion of a symmetry-conserving process which may be adequately described by the allosteric model of Monod, Wyman and Changeux⁶ (MWC) as detailed in Table 1. The data thus fit a model in which the binding of the first hapten leads to a conversion of both sites from a lower affinity conformation T into a higher affinity conformation R . It is hard to conceive such a transition in terms of local changes limited to a few contact residues only. Rather, the data suggest that hapten binding causes a change in the overall configuration of the dimer which is related to inherent structural and energetic properties of the L_2 molecule as a whole. In view of the evidence that in L_2 haptens interact with both L chains simultaneously^{1,5,16}, it is possible that the hapten participates in the inter-chain interactions that mediate the concerted transition. More specifically, we suggest that hapten binding to one of the sites between the two L chains brings about a symmetry-conserving change in the relative chain position. This transition is associated with a relatively small net free energy change, so that haptens having association constants of the magnitude found here may shift the equilibrium appreciably (see Table 1).

In analogy to the L_2 case, haptens that bind to HL also interact with both chains^{9,17,18}, sometimes enhancing their mutual cohesion¹⁹. Antigen or hapten-induced conformational transitions detected in HL by spectroscopic^{8,10} or kinetic^{12,20} methods, were actually proposed (on the basis of indirect inference) to involve changes of relative chain position^{8,19}. Our results strongly support these suggestions in showing that such changes are energetically feasible in a structural and functional homologue of an immunoglobulin. The inherent structural properties of L_2 which lead to the observed concerted conformational transition probably exist also in HL, since conformation and chain interactions in both were shown to be very similar^{1-3,17}. Direct hapten binding measurements cannot detect conformational changes in HL since it binds only one hapten per two chains (see ref. 12). Kinetic measurements, however, give information on such processes as shown by the chemical relaxation study of hapten-induced conformational changes in another mouse myeloma protein, MOPC 460 (ref. 12). In order to relate the conformational transitions in L_2 and HL, L_2 315 is now also being characterised by chemical relaxation. Preliminary results are in good agreement with

the binding data, suggesting two slowly interconverting conformations which bind the hapten differently.

The changes in relative chain position may also be expressed in a change of the domain interactions, hypothesised to be the primary event in antigen triggering of antibody effector functions through an allosteric transition¹¹. The mechanism proposed here may imply that this transition arises chiefly from mere simultaneous interaction with H and L chains, stemming from inherent structural and energetic properties of HL. It would therefore be relatively antigen-independent, which explains how different antigens, interacting with diverse sites, are able to trigger the effector functions common to all antibody molecules. In any antibody-hapten system, however, the conformational effect will depend on the detailed interactions and will not always be expected to occur.

DORON LANCET

ARIEH LICHT

ISRAEL SCHECHTER

ISRAEL PECHT

Department of Chemical Immunology,
The Weizmann Institute of Science,
Rehovot, Israel

Received 13 July; accepted 19 August 1977.

- Edmundson, A. B. *et al. Biochemistry* 13, 3816-3827 (1974).
- Edmundson, A. B., Ely, K. R., Abola, E. E., Schiffer, M. & Panagiotopoulos, N. *Biochemistry* 14, 3953-3961 (1976).
- Edmundson, A. B. *et al. in Antibodies in Human Diagnosis and Therapy* (eds Haber, E. & Krause, R. M.) (Raven, New York 1977).
- Schechter, I., Ziv, E. & Licht, A. *Biochemistry* 15, 2785-2790 (1976).
- Licht, A., Lancet, D., Schechter, I. & Pecht, I. *FEBS Lett.* 78, 211-215 (1977).
- Monod, J., Wyman, J. & Changeux, J. T. *J. molec. Biol.* 12, 88-118 (1965).
- Lancet, D. & Pecht, I. *Biochemistry* (in the press).
- Schlessinger, J., Steinberg, I. Z., Givol, D., Hochman, J. & Pecht, I. *Proc. natn. Acad. Sci. U.S.A.* 72, 2775-2779 (1975).
- Poljak, R. J., Amzel, L. M. & Phizackerley, R. P. *Prog. Biophys. molec. Biol.* 31, 67-93 (1976).
- Metzger, H. *Adv. Immun.* 18, 169-207 (1974).
- Huber, R., Deisenhofer, J., Colman, P., Matsushima, M. & Palm, W. *Nature* 264, 415-420 (1976).
- Lancet, D. & Pecht, I. *Proc. natn. Acad. Sci. U.S.A.* 73, 3549-3553 (1976).
- Adair, G. S. *J. biol. Chem.* 63, 529 (1925).
- Levitzki, A. *in Subunit Enzymes: Biochemistry and Function* (ed. Ebner, K. E.) (Marcel Dekker, New York, 1975).
- Levitzki, A. & Schlessinger, J. *Biochemistry* 13, 5214-5219 (1974).
- Stevenson, G. T. *Biochem. J.* 133, 827-836 (1973).
- Cathou, R. E., Dorrington, K. J. *in Subunits in Biological Systems*, part C (eds Timasheff, S. N. & Fasman, G. D.) (Marcel Dekker, New York, 1975).
- Padlan, E. A., Davies, D. R., Pecht, I., Givol, D. & Wright, C. *Cold Spring Harb. Symp. quant. Biol.* XLI, 627-637 (1977).
- Cathou, R. E. & Werner, T. C. *Biochemistry* 9, 3149-3155 (1970).
- Pecht, I. & Lancet, D. *in Chemical Relaxation in Molecular Biology* (eds Pecht, I. & Rigler, R.) (Springer, Heidelberg, 1977).
- Powell, M. J. D. *in Harwell Subroutine Library* (ed. Hopper, K. J.) (Subroutine VA04A Atomic energy Research Establishment, Harwell).
- Janeway, C. A., Jr, Wigzell, H. & Binz, H. *Scand. J. Immun.* 5, 993-1001 (1976).
- Krawinkel, U. *et al. Cold Spring Harb. Symp. quant. Biol.* 41, 285-294 (1977).

Two enzymes are required for strand incision in repair of alkylated DNA

THE excision-repair model for DNA repair involves the incision by an endonuclease of the phosphodiester chain adjacent or close to the damage and subsequent excision by an exonuclease of the defective nucleotide(s) (for review see ref. 1). Specific endonuclease for ultraviolet-induced damage² and for apurinic sites have been isolated associated with^{2,3}, or without an exonuclease activity (refs 4-6 and J.L. and L. Grossman, in preparation). Another class of enzymes does not incise DNA, but recognises and excises defective bases as uracil in the DNA. Belonging to this class of DNA *N*-glycosidases, is the enzyme described here which excises 3-methyladenine (3-MeAde) from alkylated DNA. This base excision repair is a preparatory step for the nucleotide excision repair described

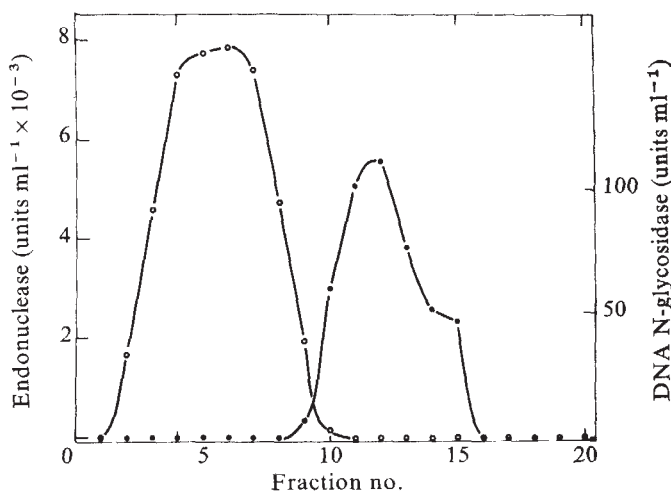


Fig. 1 Separation of DNA *N*-glycosidase from endonuclease for apurinic sites on DNA-Agarose. ●, DNA *N*-glycosidase; ○, endonuclease for apurinic sites. *M. luteus* ATCC 4698 was grown to late exponential phase¹³ washed in 1 mM EDTA 5 mM mercaptoethanol, 10% glycerol (solution A) containing 10 mM K₂PO₄ (final pH 6.5) and stored at -20 °C. Freshly grown cells (14 g), dispersed in 30 ml of washing buffer were lysed with 50 mg of lysozyme, cooled and sonicated to reduce the viscosity. The resulting solution was slowly titrated to pH 7.7 with 0.1 M KOH and centrifuged for 60 min at 30,000g. The supernatant was loaded on a DEAE-cellulose column (DE 52 Whatman) equilibrated with solution A containing 0.05 M K₂PO₄ (final pH 7.7). Elution was achieved by a linear gradient, the initial concentration was the equilibrating buffer and the limiting one, solution A, containing 0.35 M K₂PO₄ (final pH 7.7). The two activities were not separated. Active fractions were pooled, dialysed against solution A containing K₂PO₄ 0.05 M, final pH 6.7 and loaded on a P-cellulose column (P₁₁ Whatman) equilibrated with the same buffer. Elution was by a KCl linear gradient, 0-0.3 M in the loading buffer. The two enzymes were not separated. Active fractions are pooled and at this step could be stored at -20 °C with a loss of activity less than 10% in 10 months. The last step was chromatography on DNA-Agarose. P-cellulose fractions (2 ml) were diluted with 4 ml of solution A and applied to a DNA-Agarose¹⁴ column (0.4 × 8 cm) equilibrated with solution A containing 0.02 M K₂PO₄ buffer (final pH 7.2). The column was washed with 5 ml of the loading buffer, and eluted with a gradient of KCl 0-1 M in the same buffer (20 ml). Endonuclease for apurinic sites was measured by a modification of the method of Verly and Rassart³ using *E. coli* ³H DNA (3.2 × 10⁷ c.p.m. μmol⁻¹) heavily alkylated with MMS and subsequently heat depurinated. Incubation mixtures (100 μl) contained 10 μmol of K₂PO₄ buffer (pH 7.5), 4 pmol of apurinic DNA and enzyme. After incubation at 37 °C the mixture was chilled at 0 °C and 50 μl of cold calf thymus DNA (0.5 mg ml⁻¹) was added, after 2 min, 200 μl of cold 0.5 M HClO₄ was added¹⁵. After 5 min at 0 °C the precipitate was removed by a 12-min centrifugation at 6,500g and radioactivity measured in Scintigel (Roth). One unit of enzyme liberates 1 nmol of acid-soluble oligonucleotides in 30 min in the above conditions. DNA *N*-glycosidase was assayed by measuring the release of ethanol soluble 3-MeAde from a calf thymus DNA (4 mg in 2 ml) alkylated¹⁶ with ³H-dimethylsulphate (5 mCi, Amersham). Assay mixtures (100 μl) contained 10 μmol of K₂PO₄ buffer (pH 7.5), 4.58 nmol dimethylsulphate-alkylated DNA (1.5 × 10⁶ c.p.m. μmol⁻¹) and enzyme. After incubation at 37 °C, 50 μl of a solution containing 12.5 μg of calf thymus DNA and 100 μg of bovine serum albumin were added, chilled to 0 °C for 2 min and 300 μl of pre-cooled (-20 °C) ethanol are added. After an additional 5 min at 0 °C the precipitate is removed by a 12-min centrifugation (6,500g, 0 °C) and radioactivity determined. One unit of enzyme liberates 1 pmol of ethanol-soluble alkylated base in 30 min at 37 °C in the above conditions. Ethanol-soluble products are identified as 3 MeAde by chromatography on Dowex-50 (NH₄⁺ form) at pH 6.65 (ref. 17) using authentic samples of 3-MeAde(Fluka) 7-MeGua, thymine (Sigma) and 7-methyladenine (Vega-Fox) as internal standards. The chromatographic identity of ³H-3-MeAde with an authentic sample was independently confirmed using cellulose thin-layer chromatography (Eastman chromatogram)⁸.

above. It has been claimed that endonuclease II of *Escherichia coli*⁸ both excises the alkylated bases 3-MeAde and *O*⁶-methylguanine (*O*⁶-MeGua) and incises DNA at

apurinic sites. This enzyme also recognises a number of additional substrates⁹⁻¹¹. To establish the precise mechanism of incision of alkylated DNA we have purified two enzymes from *Micrococcus luteus*, a DNA *N*-glycosidase, which excises 3-MeAde, and an endonuclease for apurinic sites. We report here a reconstruction experiment showing that incision of DNA treated with carcinogenic and mutagenic compound methyl methanesulphonate (MMS) is a two-step mechanism involving the sequential action of these two enzymes.

Endonuclease for apurinic sites was assayed by the release of acid-soluble oligonucleotides³ from heavily depurinated *E. coli* ³H-DNA. DNA *N*-glycosidase activity is measured by the liberation of ethanol soluble ³H-3-MeAde (Fig. 1) from calf thymus DNA alkylated with ³H-dimethylsulphate. Supertwisted PM2-DNA was used in additional assays to detect endonuclease activity at the level of one nick per molecule. The assay measures the conversion of supertwisted molecules to nicked ones. When using superhelical DNA in the assay of the endonuclease for apurinic sites, a limited number of apurinic sites was introduced by heating at low pH (ref. 19). For DNA *N*-glycosidase few bases were alkylated by MMS²⁰.

The two enzyme activities showed, in our conditions, similar chromatographic properties on DEAE- and P-cellulose columns, but were resolved by a column of DNA-Agarose (Fig. 1). But slight contaminating traces of endonuclease acting at apurinic sites in DNA *N*-glycosidase could be detected up to fraction 12 when using a large excess of DNA *N*-glycosidase (1.1 units) for the PM2-DNA assay. No DNA *N*-glycosidase can be detected in the endonuclease for apurinic sites peak before fraction 7 using 160 units of this enzyme. Neither of the two enzymes were contaminated by nonspecific endonuclease as they did not nick supercoiled PM2-DNA (Table 1). Endonuclease for apurinic sites acts at or near an apurinic site and has no exonuclease associated (J.L. and L. Grossman, in preparation). DNA *N*-glycosidase liberates 3-MeAde from dimethylsulphate-alkylated DNA and very little, if any, of 7-methylguanine (7-MeGua) (less than 5% of the total bases released).

A reconstruction experiment involving the sequential action of the two enzymes described above should be a strong argument for a two-step mechanism for incision of alkylated DNA. Therefore, we devised the following experiment: ³H-labelled PM2-DNA was treated with MMS²⁰ to alkylate about one base per molecule as 3-MeAde²¹, this latter base being a substrate for DNA *N*-glycosidase. Identification of DNA superhelical and nicked forms was achieved by zone sedimentation or by filter assay and average number of nicks per molecule calculated. Untreated PM2-DNA carried an average of 0.06 nicks per molecule (Table 1). Enzymes used had no activity on native DNA as after incubation the average number of nicks (0.10) was not significantly increased (Table 1). This rules out a possible contamination by non-specific endonuclease. Methylation slightly nicks the DNA (Table 1). On incubation of MMS-treated DNA alone or with apurinic endonuclease or with DNA *N*-glycosidase, there was no significant change in the average number of nicks introduced per molecule of DNA, this number being respectively 0.15, 0.17, 0.16 for the three samples studied (Table 1 and Fig. 2 *a, b, c*). In sharp contrast sequential incubation of the two enzymes induced an average of 1.05 breaks per molecule (Fig. 2*d*).

It is the first time that the DNA *N*-glycosidase for alkylated DNA has been obtained devoid of endonucleolytic activity either on native DNA or on DNA-carrying apurinic sites (Fig. 1) and devoid of nicking activity on MMS-treated DNA. Similar specificity has been observed for endonuclease II (ref. 8). The specificity of the DNA *N*-glycosidase for 3-MeAde but not for 7-MeGua correlates

with the *in vivo* studies of Lawley and Warren²², showing that in *E. coli* 3-MeAde is lost much more rapidly than 7-MeGua.

Strauss *et al.*²⁰ showed that crude extracts of *M. luteus* as well as of *Bacillus subtilis* are able to introduce single-strand breaks in alkylated DNA, and that unidentified

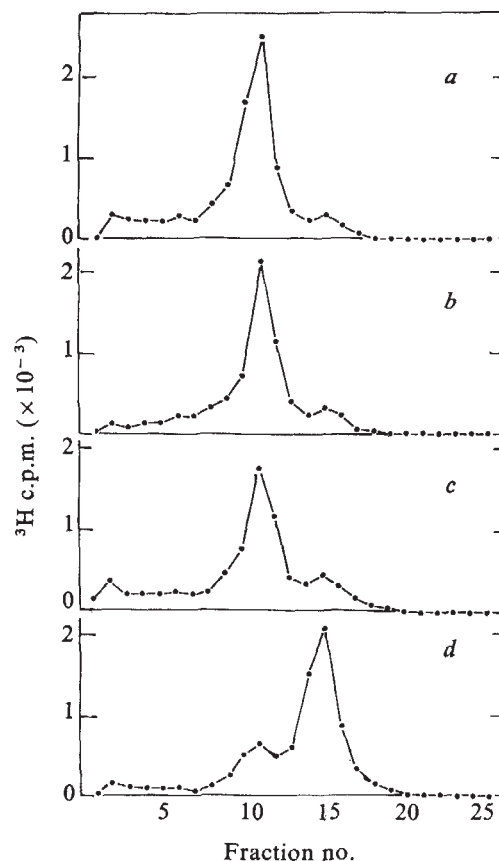


Fig. 2 Zone sedimentation profiles of PM2-DNA after incubation without enzyme (*a*) with DNA *N*-glycosidase (*b*), with endonuclease acting at apurinic sites (*c*) and with DNA *N*-glycosidase then endonuclease for apurinic sites (*d*). ³H-labelled PM2-DNA (17 nmol, 3.2×10^7 c.p.m. μmol^{-1}) was alkylated with MMS as described by Strauss *et al.*²⁰ (0.0125 M for 31 min at 37 °C), cooled to 0 °C and dialysed against 5 mM K_2PO_4 buffer (pH 7.5). This substrate is referred as MMS-DNA. The DNA *N*-glycosidase incubation mixture (150 μl) contained 26 pmol of MMS-DNA in K_2PO_4 buffer 0.05 M (pH 7.5) and 0.066 units of DNA *N*-glycosidase (fraction no. 16) incubation is for 20 min at 37 °C. To show the action of apurase the incubation mixture is as above, DNA *N*-glycosidase being replaced by 0.012 units of endonuclease for apurinic sites acting for 10 min at 37 °C. Sequential action of the two enzymes is performed by incubating 0.066 units of DNA *N*-glycosidase for 10 min at 37 °C, then adding 0.012 units of endonuclease acting at apurinic sites and incubating the mixture for an extra 10 min. For the action of the enzyme on native PM2-DNA conditions were the same, MMS-DNA being replaced by the same amount of native ³H-labelled PM2-DNA. At the end of the incubation time the samples were cooled to 0 °C and layered on the top of a 4.8-ml sucrose gradient¹⁸ (5–20% in 0.02 M Tris-HCl, 0.002 M EDTA, 0.04 M NaCl (pH 7.08) over a 0.2 ml cushion of saturated CsCl. The samples were then centrifuged (rotor SW 50-1) for 180 min at 4 °C at 43,000 r.p.m. (Spinco L2-75-B). Gradients were collected using an ISCO UA5 fraction collector and counted in Scintigel (Roth). Separation was achieved at low salt concentration which allows the best separation at neutral pH¹⁸. In these conditions, sedimentation coefficients were respectively 29S and 21S for superhelical and nicked forms¹⁸. Centrifugation in alkaline conditions although allowing even better separations¹² cannot be used as they nick DNA at apurinic sites¹⁹. But alkaline denaturation can be performed in the filter assay as time of contact is short (30 s) and without adverse effect (Table 1).

Table 1 Average no. of nicks introduced by various treatments of PM2-DNA

DNA treatment	Average no. of nicks
None (untreated DNA)	0.06
DNA + MMS + dialysis (MMS-DNA)	0.10
MMS-DNA + incubation	0.15
MMS-DNA + DNA <i>N</i> -glycosidase	0.17
MMS-DNA + endonuclease for apurinic sites	0.16
MMS-DNA + DNA <i>N</i> -glycosidase + endonuclease for apurinic sites	1.05
Untreated DNA + DNA <i>N</i> -glycosidase + endonuclease for apurinic sites	0.10

PM2-DNA was alkylated by MMS, dialysed and treated with the various enzymes as duplicates of experiments described in Fig. 2. After incubation, samples were cooled to 0 °C. The DNA was immediately denatured by alkaline treatment, renatured and filtered through nitrate cellulose filter as described previously¹². Single-stranded DNA arising from nicked molecules was retained on the filter, allowing calculation of the amount of supertwisted form in the incubation mixture. The number of breaks (n) per molecule was derived from Poisson's law, $n = -\ln e$, e being the fraction of remaining supertwisted molecules²³.

methylated sites are not substrate for the enzyme(s). These early experiments had not been devised to ascertain if incision of DNA occurred through a single or a two-step mechanism, however. Goldthwait's group showed that a purified although non-homogeneous preparation of *E. coli* endonuclease II, is active on a wide range of substrates⁸⁻¹¹—it nicks depurinated DNA, releases 3-MeAde and *O*⁶-methylguanine from DNA reacted with *N*-methyl-*N*-nitrosourea⁸, and also nicks unmodified DNA at high concentrations⁹. Verly and Rassart³ isolated an endonuclease in *E. coli* acting at an apurinic site which was obtained homogeneous and shown to be inactive on alkylated DNA. This endonuclease was found to be exonuclease III (ref. 24), a single molecule carrying the two activities assayed independently with different substrates. Kirtikar *et al.*²⁵ isolated from *E. coli* an apurinic acid endonuclease apparently the same as that isolated by Verly and Rassart³ but they consider it not to be exonuclease III. In *E. coli*, another endonuclease acting at apurinic sites and lacking exonuclease activity (endonuclease IV) was obtained homogeneous and shown to be inactive on alkylated DNA⁵.

The model of excision repair of alkylated DNA best supported by our data and the published data on the *M. luteus* system is shown in Fig. 3; part of this scheme was proposed as a hypothetical model by Lindahl⁷. On action of a variety of alkylating agents on DNA (Fig. 3a) at least *N*-7-guanine and *N*-3-adenine became alkylated²¹ (Fig. 3b). 3-MeAde but not 7-MeGua is recognised by the DNA *N*-glycosidase. It cleaves the *N*₉ C'1 bond between the base and the deoxyribose, yielding a free 3-MeAde and an apurinic site (Fig. 3c). In turn, the apurinic site is recognised by the endonuclease for apurinic site which nicks the DNA (Fig. 3d). It should be recalled that *M. luteus* endonuclease for apurinic site has no exonuclease activity. This could be a disadvantage as DNA ligase could act before the removal of the damaged nucleotide and thus produce an abortive repair. It is possible, however, that DNA binding proteins could prevent this abortive repair (Fig. 3e) as this has been already suggested in the repair of ultraviolet-induced damage¹. The correxonuclease described by Kushner *et al.*²⁶ which is able to excise thymine dimers would be a good candidate to excise apurinic sites yielding a gap (Fig. 3e). The gap thus formed could be filled by a DNA polymerase²⁷ (Fig. 3f) and the continuity of the phosphodiester chain restored by a polynucleotide ligase (Fig. 3g). This model has recently received experimental support in the repair

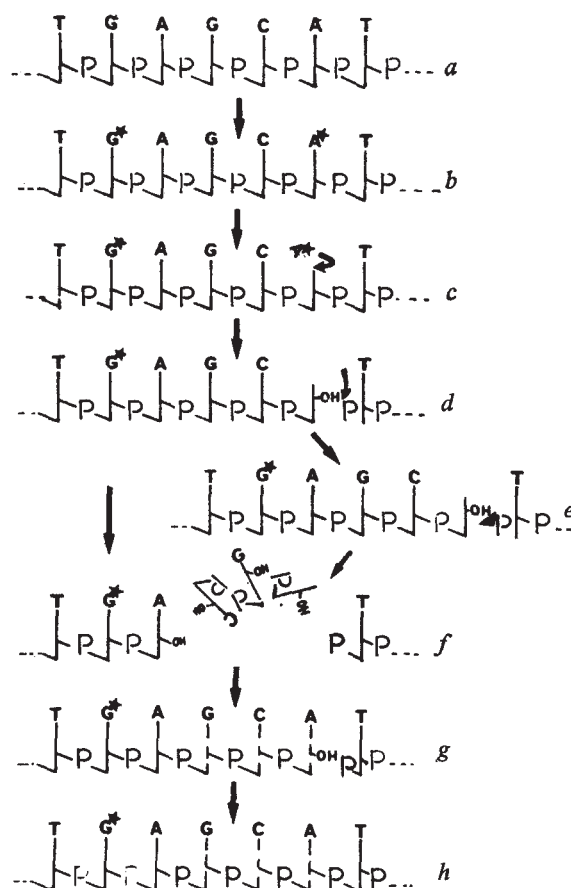


Fig. 3 Model for excision repair of alkylated DNA or uracil-containing DNA. *a*, Unmodified DNA; *b*, alkylation (MMS) giving 3-MeAde (A*) and 7-MeGua (G*); *c*, specific release of free 3-MeAde yielding an apurinic site (DNA *N*-glycosidase); *d*, incision endonuclease for apurinic sites; *e*, a binding protein prevents abortive resealing; *f*, excision (correxonuclease)—this step could also occur directly after (*d*); *g*, repair replication (DNA polymerase); *h*, resealing (DNA ligase). This model holds true for repair of uracil-containing DNA²⁸ (A* being replaced by U).

of DNA containing uracil where nicking is achieved by sequential action of uracil DNA *N*-glycosidase and endonuclease for apurinic sites²⁸. Endonucleases acting at apurinic sites have been described in calf thymus⁴, HeLa cells²³ and human placenta²⁹. A human binding protein for ultraviolet irradiated DNA³⁰ and a correxonuclease³¹ are described in human placenta. Therefore, our model could also be true for repair in mammalian cells of alkylated DNA, even if so far no mammalian DNA *N*-glycosidase has been described.

In animals treated with alkylating carcinogens (see review 32), certain organs such as the liver seem to be much more resistant to tumourigenesis than the brain or the kidneys. Appearance of tumours is correlated with the persistence of alkylated bases. This suggests that chemical carcinogenesis could be related to defective repair mechanisms such as inefficient cellular DNA *N*-glycosidase.

We thank Drs S. Ljungquist and T. Lindahl for preprints. These studies were supported by INSERM and CNRS and Pierre Philippe Foundation.

JACQUES LAVAL

Laboratoire Associé n 147
CNRS et Unité 147 INSERM,
Institut Gustave-Roussy, 94800 Villejuif, France

Received 4 July; accepted 9 September 1977.

1. Grossman, L., Braun, A., Feldberg, R. & Mahler, I. *A. Rev. Biochem.* **44**, 19–43 (1975).
2. Hadi, S. M. & Goldthwait, D. A. *Biochemistry* **10**, 4986–4994 (1971).
3. Verly, W. G. & Rassart, E. *J. biol. Chem.* **250**, 8214–8219 (1975).
4. Ljungquist, S., Nyberg, B. & Lindahl, T. *FEBS Lett.* **57**, 169–171 (1975).
5. Ljungquist, S. *J. biol. Chem.* **252**, 2808–2814 (1977).
6. Thibodeau, L. & Verly, W. G. *FEBS Lett.* **69**, 183–185 (1976).
7. Lindahl, T. *Nature* **259**, 64–66 (1976).
8. Kirtikar, D. M. & Goldthwait, D. A. *Proc. natn. Acad. Sci. U.S.A.* **71**, 2022–2026 (1974).
9. Hadi, S. M., Kirtikar, D. & Goldthwait, D. A. *Biochemistry* **12**, 2747–2754 (1973).
10. Kirtikar, D. M., Slaughter, J. & Goldthwait, D. A. *Biochemistry* **14**, 1235–1243 (1975).
11. Kirtikar, D. M., Dipple, A. & Goldthwait, D. A. *Biochemistry* **14**, 5548–5553 (1975).
12. Laval, F. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4965–4969 (1974).
13. Kaplan, J. C., Kushner, S. R. & Grossman, L. *Proc. natn. Acad. Sci. U.S.A.* **63**, 144–151 (1969).
14. Arndt-Jovin, D. J., Jovin, T. M., Bähr, W., Frischauf, A. M. & Marquardt, M. *Eur. J. Biochem.* **54**, 411–418 (1975).
15. Laval, J. & Paoletti, C. *Biochemistry* **11**, 3596–3603 (1972).
16. Uhlenhopp, E. L. & Krasna, A. I. *Biochemistry* **10**, 3290–3295 (1971).
17. Lawley, P. D. & Shah, S. A. *Biochem. J.* **128**, 117–132 (1972).
18. Bottger, M. & Kuhn, W. *Biochim. biophys. Acta* **254**, 407–411 (1971).
19. Lindahl, T. & Anderson, A. *Biochemistry* **11**, 3618–3623 (1972).
20. Strauss, B., Coyle, M. & Robbins, M. *Cold Spring Harb. Symp. quant. Biol.* **33**, 277–287 (1968).
21. Lawley, P. D. & Shah, S. A. *Chem.-Biol. Interact.* **5**, 286–288 (1972).
22. Lawley, P. D. & Warren, W. *Chem.-Biol. Interact.* **12**, 211–220 (1976).
23. Teebor, G. W. & Duker, N. J. *Nature* **258**, 544–547 (1975).
24. Weiss, B. *J. biol. Chem.* **251**, 1896–1901 (1976).
25. Kirtikar, D. M., Cathcart, G. R. & Goldthwait, D. A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4324–4328 (1976).
26. Kushner, S. R., Kaplan, J. C., Ono, H. & Grossman, L. *Biochemistry* **10**, 3325–3334 (1971).
27. Hamilton, L., Mahler, I. & Grossman, L. *Biochemistry* **13**, 1886–1896 (1974).
28. Lindahl, T., Ljungquist, S., Siegert, W., Nyberg, B. & Sperens, B. *J. biol. Chem.* (in the press).
29. Linsley, W. S., Penhoet, E. E. & Linn, S. *J. biol. Chem.* **252**, 1235–1242 (1977).
30. Feldberg, R. S. & Grossman, L. *Biochemistry* **15**, 2402–2408 (1976).
31. Doniger, J. & Grossman, L. *J. biol. Chem.* **251**, 4579–4587 (1976).
32. Pegg, A. E. *Adv. Cancer Res.* **25**, 195–269 (1977).

Mass spectrometric determination of molecular formulas for membrane-modifying antibiotics

THE peptide antibiotics alamethicin and suzukacillin, together with gramicidin A are of considerable biological interest as they facilitate ion transport across membranes by a mechanism involving pore or channel formation in the membranes^{1–3}. In discussing their mode of action it is, of course, necessary to consider the antibiotics' structures and such discussions have been hampered, especially for alamethicin^{1,4–7}, by incorrect structures originally

proposed for the antibiotics. Thus, the earlier structures proposed for valine-gramicidin A and isoleucine-gramicidin A^{8,9} were revised to **1** and **2** (Fig. 1) following the detection of N-terminal formyl and C-terminal ethanolamine groups¹⁰, and earlier structures^{11,12} proposed for alamethicins I and II have recently been revised to **3** and **4** (ref. 13) following the discovery of N-terminal acetyl and C-terminal phenylalaninol groups^{14,15}. (The partial structure proposed for suzukacillin¹⁶ will probably also be revised in view of the newly reported structure for alamethicin¹³.) Much of the difficulty in assigning structures to these antibiotics has derived from an inability to determine accurate molecular formulas. We have until recently been hampered by the same problem in our studies of the structures of antiamoebin I (**5**) (ref. 17) and emerimicins III and IV (**6** and **7**) (ref. 18), new peptaibophol antibiotics [peptide antibiotics containing several moles of α -aminoisobutyric acid (Aib) and a C-terminal phenylalaninol (Phol) unit] which have been shown to have similar membrane-modifying properties (P. Mueller, personal communication). We have developed, and describe here, a new technique, a modification of field desorption mass spectrometry^{19,20} involving cation exchange, which has allowed us to assign molecular formulas directly to these peptide antibiotics and their derivatives in the 1,500–2,100 molecular weight range.

The cation exchange technique involves adding, systematically, a series of alkali metal salts to solutions of the compounds whose FD mass spectra are being determined, and observing the resulting shifts in the molecular ion region due to cationation. The technique was developed during our extensive studies on the membrane-modifying peptide antibiotic antiamoebin I, whose structure has now been shown²¹ to be **5**. We found it difficult to assign the antibiotic a definitive molecular formula, either from its constituent amino acids or from the highest mass ion in the FD mass spectrum of its triacetate, at m/e 1,818 (peak matched against the m/e 1,790 peak ($C_{35}H_{15}F_{60}N_3O_5P_3$) from hexakis (2,2,3,3,4,4,5,5,6,6,7,7, - dodecafluoro - 1 - heptoxycyclotriphosphazene²²), which shifted to m/e 1,860 and 1,902 in the FD mass spectra of the tripropionate and tributrylate, respectively.

The masses for the triesters would correspond to a molecular weight of 1,692 for antiamoebin I, which did not agree with any of the possibilities under consideration¹⁷. But, m/e 1,692 would fit for a sodium adduct of one of the considered potential molecular weights, m/e 1,669 ($C_{82}H_{127}N_{17}O_{20}$). To test this hypothesis

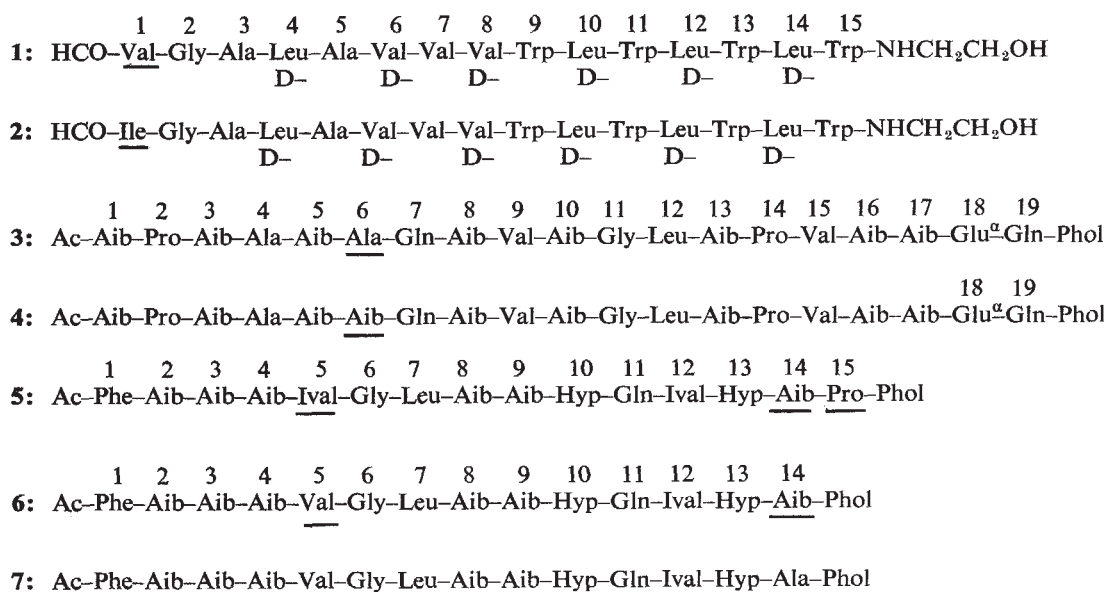


Fig. 1 **1**, Valine-gramicidin A; **2**, isoleucine-gramicidin A; **3**, alamethicin I; **4**, alamethicin II; **5**, antiamoebin I; **6**, emerimicin III; **7**, emerimicin IV. All optically active amino acids and phenylalaninol units shown have the L configuration except those marked D. Aib, α -aminoisobutyric acid; Phol, phenylalaninol; Ac, acetate.

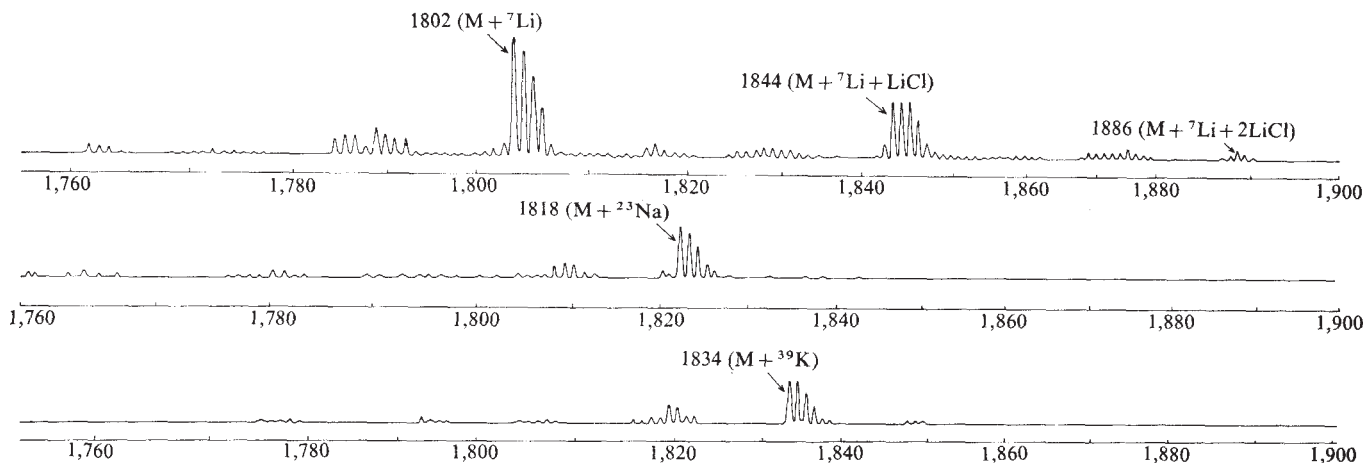


Fig. 2 Field desorption mass spectra of antimoebin I triacetate in the molecular ion region: *a*, with added lithium chloride (26 mA); *b*, with added sodium chloride (23 mA); *c*, with added potassium chloride (21 mA). Peaks due to a small amount of a homologue, antimoebin III, can be observed 14 AMU below the major peaks.

antimoebin triacetate was treated with a saturated methanolic solution of potassium chloride and a 50% saturated solution of lithium chloride. The triacetate's m/e 1,818 peak ($M + Na$; Fig. 2, *b*) was displaced to m/e 1,834 ($M + K$; Fig. 2, *c*) and m/e 1,802 ($M + Li$; Fig. 2, *a*), respectively. In the case of lithium chloride, peaks at higher masses corresponding to ($M + Li + LiCl$) and ($M + Li + 2LiCl$) were also observed (Fig. 2, *a*), at 1,844 and 1,886. (Addition of rubidium chloride can be even more helpful, since there are two rubidium isotopes (^{85}Rb and ^{87}Rb ; isotopic abundance 2.51:1). In the case of antimoebin I triacetate these gave $M + Rb$ ions at m/e 1,880 and 1,882.)

The intensities of molecular ion peaks were vastly enhanced by this treatment and enormously more durable. (Lithium salts have been added previously, using the FD technique at low resolution with known compounds, to adenosine (molecular weight (MW) 267)²³ and to sucrose (MW 342) and loriglossin (MW 742) (ref. 24). The triacetate's m/e 1,818 peak was matched against the peak at m/e 1,720.9514 ($C_{34}H_{18}F_{56}N_3O_6P_3$) in the FD spectrum of bis (dodecafluoroheptoxy) tetrakis (octafluoropentoxy)-cyclotriphosphazene²² at high resolution to give the value 1,818.9596, in good agreement with that calculated for the expected $C_{88}H_{133}N_{17}O_{23}Na$.

The technique works equally well on antimoebin itself. Adding sodium, potassium and lithium chloride solutions to antimoebin I gave strong FD peaks at m/e 1,692 ($M + Na$), 1,708 ($M + K$) and 1,676 ($M + Li$), while addition of sodium chloride solutions to the tripionate and tributrylate greatly enhanced the peaks at m/e 1,860 and 1,902 ($M + Na$).

The cation exchange technique shows the molecular formula of the membrane-modifying emerimicin IV (6) (ref. 25) to be $C_{77}H_{120}N_{16}O_{19}$ (MW 1,572) by an ($M + Na$) ion at m/e 1,595.8930 and an ($M + K$) ion at m/e 1,611, while its triacetate gave an ($M + Na$) ion at m/e 1,721 and an ($M + K$) ion at m/e 1,737. Emerimicin III (7) (ref. 25), also a membrane-modifier, gave an ($M + Na$) peak at m/e 1,581.8655.

Alamethicin is a mixture of two antibiotics, alamethicins I and II (3 and 4) (ref. 13), each containing a free carboxyl group, and its cation exchange FD mass spectra are more complex. The mixture gave peaks for alamethicin I ($C_{92}H_{150}N_{22}O_{25}$) at m/e 1,984 ($M + Na - H$) and 2,007 ($M + 2Na - H$) and corresponding peaks for alamethicin II ($C_{93}H_{152}N_{22}O_{25}$) at m/e 1,998 and 2,021 in the presence of sodium chloride. Similar peaks were found at m/e 2,000 ($M + K - H$), 2,014 ($M + K - H$), 2,039 ($M + 2K - H$) and 2,053 ($M + 2K - H$) in the presence of potassium chloride, while alamethicin acetate gave peaks at m/e 2,026 ($M + Na - H$), 2,040 ($M + Na - H$), 2,049 ($M + 2Na - H$) and 2,063 ($M + 2Na - H$) in the presence of sodium chloride.

Finally, sodium chloride-treated gramicidin A gave FD peaks at m/e 1,903 ($M + Na$ for valine-gramicidin A, $C_{99}H_{140}N_{20}O_{17}$, 1)

and at 1,917 (the same peak for isoleucine-gramicidin A, $C_{100}H_{142}N_{20}O_{17}$, 2), while treatment with potassium chloride shifted these peaks to m/e 1,919 and 1,933 ($M + K$ for valine- and isoleucine-gramicidins A, respectively).

We are now extending our studies to establish the scope of the cation exchange procedure in assigning molecular weights to additional antibiotics and to compounds of other structural classes of high molecular weight or low volatility.

This work was supported by the US NCI and the National Institute of Allergy and Infectious Diseases.

K. L. RINEHART, JR
J. C. COOK, JR
H. MENG
K. L. OLSON
R. C. PANDEY

School of Chemical Sciences,
University of Illinois,
Urbana, Illinois 61801

Received 4 April; accepted 11 July 1977.

- Haydon, D. A. & Hladky, S. B. *Q. Rev. Biophys.* **5**, 187-282 (1972).
- Mueller, P. & Rudin, D. O. *Nature* **217**, 713-719 (1968).
- Boheim, G. *et al. Biochim. biophys. Acta* **433**, 182-199 (1976).
- Boheim, G. *J. Membr. Biol.* **19**, 277-305 (1974).
- Cherry, R. J., Chapman, D. & Graham, D. E. *J. Membr. Biol.* **7**, 325-344 (1972).
- McMullen, A. I., Marlborough, D. I. & Bayley, P. M. *FEBS Lett.* **16**, 278-283 (1971).
- Jung, G., Dubischar, N. & Leibfritz, D. *Eur. J. Biochem.* **54**, 395-409 (1975).
- Synge, R. L. M. *Biochem. J.* **39**, 355-362 (1945).
- Synge, R. L. M. *Cold Spring Harb. Symp. quant. Biol.* **14**, 191-198 (1950).
- Sarges, R. & Witkop, B. *J. Am. chem. Soc.* **86**, 1861-1862 (1964); **86**, 1862-1863 (1964); **87**, 2011-2020 (1965).
- Payne, J. W., Jakes, R. & Hartley, B. S. *Biochem. J.* **117**, 767-776 (1970).
- Ovchinnikov, Y. A., Kiryushkin, A. A. & Kozhevnikova, I. V. *Gen. Chem. U.S.S.R.* **41**, 2105-2116 (1971).
- Pandey, R. C., Cook, J. C., Jr & Rinehart, K. L., Jr *J. Am. chem. Soc.* **99**, (in the press).
- Pandey, R. C., Cook, J. C., Jr & Rinehart, K. L., Jr *23rd Annual Conference on Mass Spectrometry and Allied Topics*, Houston, Texas, 25-30 May 1975.
- Martin, D. R. & Williams, R. J. P. *Biochem. J.* **153**, 181-190 (1976).
- Jung, G. *et al. Biochim. biophys. Acta* **433**, 164-181 (1976).
- Pandey, R. C., Rinehart, K. L., Jr, Meng, H., Deshmukh, P. V. & Millington, D. S. *9th International Symposium on the Chemistry of Natural Products*, IUPAC, Ottawa, Canada, 24-28 June, 1974, Abstr. 1B.
- Pandey, R. C., Rinehart, K. L., Jr, & Argoudelis, A. D. *14th Interscience Conference on Antimicrobial Agents and Chemotherapy*, San Francisco, California, 11-13 September 1974, Abstr. 302.
- Beckey, H. D. in *Biomedical Applications of Mass Spectrometry* (ed. Waller, G. R.) 795-816 (Wiley-Interscience, New York, 1973).
- Rinehart, K. L., Jr, Cook, J. C., Jr, Maurer, K. L. & Rapp, U. J. *Antibiot.* **27**, 1-13 (1974).
- Pandey, R. C., Meng, H., Cook, J. C., Jr & Rinehart, K. L., Jr *J. Am. chem. Soc.* **99**, 5203-5205 (1977).
- Olson, K. L., Rinehart, K. L., Jr & Cook, J. C., Jr *Biomed. Mass Spectrom.* **4**, (in the press).
- Röllgen, F. W. & Schulten, H. R. *Org. Mass Spectrom.* **10**, 660-668 (1975).
- Veith, H. J. *Angew. Chem. Int. Ed. Engl.* **15**, 696 (1976).
- Pandey, R. C., Cook, J. C., Jr & Rinehart, K. L., Jr *J. Am. chem. Soc.* **99**, 5205-5206 (1977).

New rapid gel sequencing method for RNA

METHODS for sequencing nucleic acids have improved significantly in the last few years due to the development of rapid 'read-off' gel methods. The first method to be developed was the 'plus and minus' copying method for DNA¹ and an analogous copying

method for RNA now exists². A different and more direct method for sequencing DNA has been developed³, but no direct method is available for RNA. We have developed such a method and here we describe its use in establishing the sequence of tyrosine tRNA of *Bacillus stearothermophilus*. This sequence is especially interesting as the tyrosine tRNA and its specific synthetase in *B. stearothermophilus* offer an attractive experimental system for studying the specific interaction of tRNA and synthetases, which is now possible because of progress in studying the three-dimensional structure of the protein by X-ray crystallography⁴.

Classical radioactive sequencing methods⁵ applied to uniformly ³²P-labelled tRNA were used initially to sequence the products of complete T₁ or pancreatic RNase digestion (Table 1). Only brief reference to this work is made here as an extensive report will be published elsewhere (R.S.B. in preparation).

Overlapping of the products in Table 1 was achieved by the development of the new sequencing procedure, which is similar in principle to that used by Maxam and Gilbert for DNA³ except that enzymatic rather than chemical methods were used to degrade the RNA and to define the base-specific point of cleavage. Thus the design of the experiment was to label the RNA at its 5' terminus with ³²P-phosphate and partially digest separate aliquots with either T₁ RNase (G-specific), U₂ RNase (A-specific in partial conditions), pancreatic RNase (C + U-specific), or RNase I from *Physarum polycephalum*⁶. Phy I RNase is less well known than the other enzymes and it emerged^{7,13} that its base specificity was such that it could cleave at all residues except C. Hence we refer to its activity as minus C (–C). The digests were then fractionated in adjacent slots of a 20% denaturing acrylamide gel. The base specificity of these four enzymes should provide enough information to allow the sequence to be 'read-off' from the resultant autoradiograph.

Figure 1 shows the purification of ³²P-5' end-labelled tRNA using γ-³²P-ATP and T₄ polynucleotide kinase by acrylamide gel electrophoresis. In addition to intact tRNA^{Tyr} (band 1), many degradation products were present in lower yield. Bands 1 and 2 were selected for sequence analysis. Figure 2 shows the sequencing method applied to the 5' end-labelled fragments of tRNA^{Tyr} (Fig. 1, band 2) which was identified as deriving from residues 38–85 in the finally deduced sequence (Fig. 4). The enzymatic degradations are identified as G, A, C + U and –C (performed in duplicate and applied at separate times so as to maximise resolution of both fast and slow moving bands). A fifth slot (hot formamide degradation) giving a ladder (L) of every possible intermediate was also included. Hot formamide degradation (P. H. Hamlyn, personal communication) is preferable to partial alkaline hydrolysis for this purpose as it gives oligonucleotides with the same 2', 3'-cyclic phosphate end group as in the enzymatic degradations. Partial alkaline hydrolysis with strong base, on the other hand, gave both

2', 3'-cyclic phosphate and a mixture of 2' and 3' phosphate end groups which separated from one another and confused the ladder in the size range of 1–10 on the acrylamide gels.

The sequence may be read from residues 38–83 with the single ambiguity X63, due to the presence of a minor base (see below). G residues are identified without difficulty from the highly specific G slot. 'A' residues are usually present in the A slot as strong bands, although also present are some G and pyrimidine bands. The latter are clearly distinguishable by their low yield and by the presence in the G or C + U slots. The pyrimidine residues are more difficult to identify than the purines, although the C + U and –C slots can be 'read' together to give the correct sequence. C residues are identified by their presence (usually as strong bands) in the C + U slot and by their absence (or low yield) in the –C slot. U residues are identified by the presence of a band in both C + U and –C slots. Usually U residues are stronger in the –C slot than the C + U slot. In this particular experiment the presence of a band, C65, in the –C slot (as well as in the G and A slots) is misleading as it was also present in a control line (not shown) and derived by nonspecific degradation of the RNA during its purification and storage. The sequence overlapped and was in complete agreement with the data in Table 1, apart from the identification of modified bases and the last two residues at the 3'-end. X63 was clearly the T residue of the oligonucleotide t9 (Table 1) and was presumably resistant to digestions in the C + U and –C slots in the partial conditions used. Both Ψ residues (residues 40 and 64) behaved like U residues, and the hyper-modified ms²i⁶A (residue 38) was only a weak band in the A slot.

Figure 3 shows the sequence analysis of intact tyrosine tRNA (Fig. 1, band 1). In this case the sequencing was more difficult than with the fragment as some degradation products were absent or in low yield. This presumably reflected the inability of the enzymes to attack helical regions in the structure. Preliminary experiments with brewers yeast tRNA^{Phe} had shown that a more even intensity of the bands was present in the G and C + U line if 70% dimethyl sulphoxide (DMSO)⁹ at 37 °C was used to denature the tRNA. Unfortunately DMSO could not be used for studies on tRNA^{Tyr} because of a 'band doubling' phenomenon on the gels, which seemed to be caused by oxidation of the s⁴U residue not present in tRNA^{Phe}. Further studies showed that 70% formamide could be used as an alternative but was not quite as effective as the DMSO. Thus Fig. 3 includes several variations in the conditions of enzymatic degradation: 'C + U, condition 3' was a reaction in formamide and was a slight improvement over the other two aqueous digestions; '–C, conditions 1 and 2' and 'C + U, conditions 1 and 2' were variations in the amount of enzyme or time of digestion in a further attempt to obtain as even a spectrum of products as possible. The sequence can, nevertheless, be read from Fig. 3 from residues 2 to 34 with only four ambiguities—at

Table 1 End products of T₁ and pancreatic RNase digestions of tRNA^{Tyr}

T ₁ end products	Pancreatic end products
t1 pG	p1 C
t2 G + G >	p2 U + ψ
t3 C – G	p3 A – C
t4 A – G	p4 Q – U
t5 U – Gm – G	p5 G – C
t6 A – A – G	p6 G – U
t7 s ⁴ U – A – G	p7 A – C – G
t8 U – U – C – G	p8 Gm – G – C
t9 T – ψ – C – G	p9 G – G – C
t10 A – A – U – C – C – G	p10 G – G – T
t11 C – U – A – A – m ¹ A – C – G	p11 A – A – m ¹ A – C
t12 C – U – C – C – C – U – U – U – G	p12 G – G – A – C
t13 A – C – U – Q – U – A – ms ² i ⁶ A – A – ψ – C – C – G	p13 G – A – A – U
t14 U – C – C – C – C – C – U – C – C – A – C – C – AOH	p14 G – G – G – U
	p15 A – ms ² i ⁶ A – A – ψ
	p16 G – A – A – G – U
	p17 pG – G – A – G – G – G – G – s ⁴ U

Uniformly ³²P-labelled tyrosine tRNA (*B. stearothermophilus*, NCIB 8924) was purified by acrylamide gel electrophoresis (R.S.B., in preparation) and fingerprints of T₁ and pancreatic RNase digestions were prepared using two-dimensional electrophoresis (cellulose acetate at pH 3.5 followed by DEAE-cellulose ionophoresis using 7% formic acid). Oligonucleotides were sequenced by standard methods⁵. Minor bases were identified by their chromatographic properties on thin layers⁸.

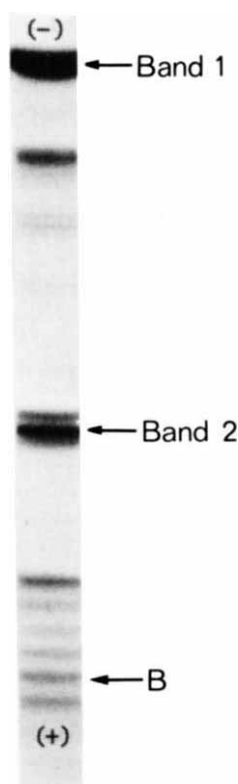


Fig. 1 15% acrylamide gel electrophoresis in 7 M urea of 5'-³²P-labelled tRNA^{Tyr}. B marks the xylene cyanol FF dye marker. Bands 1 and 2 were eluted for sequencing. 60 µg tRNA^{Tyr} was incubated in 40 µl solution (pH 8.0) containing 10 mM Tris-HCl, 0.1 mM EDTA, 10 µg bacterial alkaline phosphatase (Sigma) at 37 °C for 1 h. 10 µl nitrilotriacetic acid (40 mM) was added and the solution dried down by heating (100 °C in an open tube for about 3 min). The residue was dissolved in 100 µl solution (pH 9.5) containing 50 mM Tris-HCl, 10 mM MgCl₂, 5 mM dithiothreitol, 40 units T₄ polynucleotide phosphokinase (P.L. Biochemicals) and 4.5 µM γ-³²P-ATP (210 Ci mmol⁻¹) (Radiochemical Centre, Amersham) and the mixture was incubated at 37 °C for 15 min. An equal volume of formamide-dye mixture² was added, the mixture heated at 100 °C for 2 min and loaded on a 15% acrylamide-7 M urea gel. Electrophoresis was at 600 V for 12 h. Bands indicated (1 and 2) were eluted at 37 °C for 2 h using 0.5 M ammonium acetate (pH 7.0), 0.1 mM EDTA containing 0.5% SDS and carrier tRNA. After ethanol precipitation (twice overnight at -20 °C) the RNA samples were dissolved in distilled water and stored at -20 °C in 1 µg µl⁻¹ carrier tRNA concentration.

X17, X23, Py33 and Py34. The unusually wide spacing between positions 16 and 18 indicated a missing band (X17) in all slots, including the 'ladder'. From the sequence read from residues 15 to 18 on the gel, this missing band is clearly identified as corresponding to the 2'-*O*-methylguanosine of t5, U-Gm-G (Table 1). X23 clearly corresponds to m¹A in oligonucleotide t11. This m¹A is known to be almost completely resistant to U₂ RNase¹⁰—hence it fails to appear in the A line. Residues 33 and 34 could only be identified as pyrimidines. Beyond residue 34 the sequence could not be read, mainly because of 'missing' bands caused by the secondary structure of the tRNA^{Tyr}.

The sequences derived from the sequencing gels (Figs 2 and 3) did not overlap with one another, but they could be overlapped with the long T₁ RNase fragment (t13 of Table 1) to deduce the complete sequence (Fig. 4).

The rapid gel sequencing method described here is probably capable of improvement. With the present technology it may not yet be absolutely accurate without confirmatory data from more classical methods. There is no problem in assigning purine residues (G and A), but pyrimidines (C and U) are sometimes difficult. A correct sequence was read for all residues, however, including pyrimidines, in the sequencing gels presented here for tyrosine tRNA. Additional information which may help in the assignment of pyrimidine residues is available from the detailed kinetic studies

of Pilly *et al.*⁷ on model dinucleotides with the *Phy* I RNase. Briefly, the nature of the base on the 3' side (N) of the cleavage has differential effects on the hydrolysis of the CpN and UpN bonds. Also, C and U residues have differential effects on the rate of hydrolysis of the internucleotide bond, depending on the base (G or A) to their 5' side. Although this detailed information was not used to read the gels reported here, it might be important in other applications of the method.

The method would be improved if other enzymes were available to distinguish C and U residues that had a simpler specificity than *Phy* I RNase. Human RNase C¹¹ was tested but did not distinguish C and U residues. In theory the water-soluble carbodiimide (*N*-cyclohexyl-*N'*-(methylmorpholinium)-ethyl carbodiimide-*p*-toluene sulphonate), which blocks the action of pancreatic RNase at U residues, should be useful. But this would depend ideally on chemical modification of the RNA by this reagent approaching 100%. In addition de-protection of the RNA should be achieved without concomitant nonspecific breakdown of internucleotide bonds. Both these points pose technical difficulties that have not yet been overcome. A further factor that would undoubtedly improve the sequencing method would be to define fully denaturing conditions for the RNA which were still compatible with the activity of the RNases. This should even-out the intensity of the bands on the gels and thus make 'reading' easier and more reliable. We observed that U₂ and *Phy* I RNases were more sensitive to the

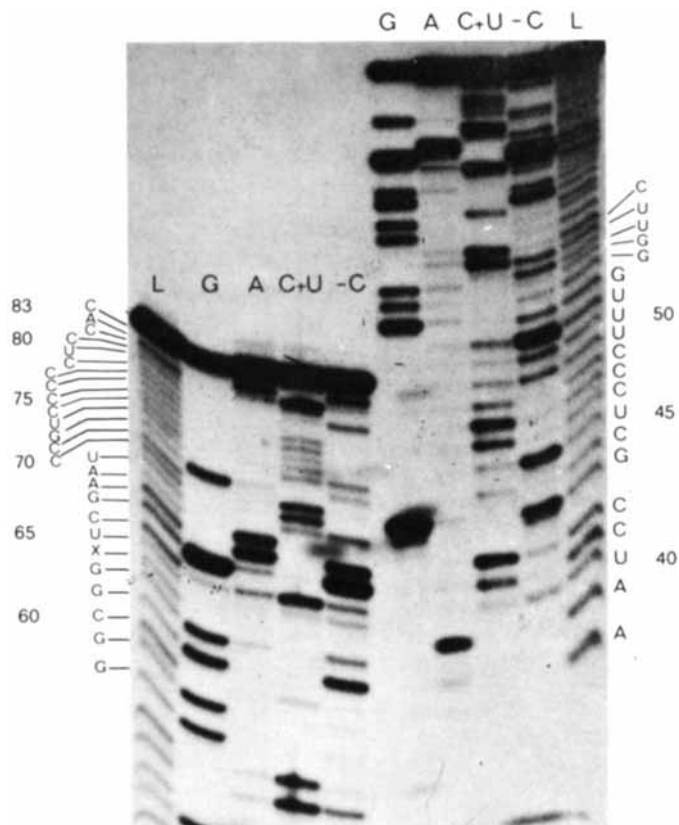


Fig. 2 Sequence analysis of 'fragment' of tRNA^{Tyr}. L shows the 'ladder' obtained by heating 1 µg RNA in 'AnalAv' formamide (10 µl, 100 °C for 30 min). Enzymatic hydrolyses were carried out using 1 µg RNA in 10 µl reaction volume as follows: G, 0.2 units (U) T₁ RNase, 0.1 M Tris-HCl, 10 mM EDTA, pH 7.5, 0 °C, 30 min; A, 0.1 U U₂ RNase, 50 mM sodium acetate, 2 mM EDTA, pH 4.5, 0 °C, 30 min; C + U, 0.1 ng pancreatic RNase, pH 7.5 buffer (as for T₁ RNase), 0 °C, 30 min; -, C, 1 µl 7 U ml⁻¹ *Phy I* RNase in 40% glycerol was added to 1 µg RNA in 9 µl 10 mM sodium acetate, 1 mM EDTA buffer (pH 5.0), room temperature, 15 min. Reactions were stopped by adding 15 µl formamide-dye mixture and heating at 100 °C for 1 min. Two separate loadings of each sample at different times were used. Gel electrophoresis was carried out on a 20% acrylamide-7 M urea gel³ using about 30,000 d.p.m. per reaction. A sensitive 'preprofing, screen' procedure was used for radioautography (R. Laskey, personal communication).

partially denaturing conditions used than were T_1 or pancreatic RNases. Nevertheless tRNA is a particularly difficult example with its 'tight' structure, and many RNA molecules to which the method is applicable would have far less secondary structure.

The technique described here compares favourably in its rapidity and sensitivity (less than $0.1 \mu\text{Ci}$ RNA per gel is needed) with the plus and minus method for RNA²; furthermore it is independent of any requirement for oligonucleotide primers². Although we do not claim that the method is 100% accurate, it is promising enough to apply to other RNA molecules. There is no doubt that it is the easiest way of establishing a tentative new sequence, which could be confirmed by more classical methods if necessary. The method in principle is also applicable to 3' terminally-labelled RNA. Preliminary results have shown that ^{32}P 3'-labelled tRNA^{Phe} (labelled using $\alpha\text{-}^{32}\text{P}$ -ATP and the enzyme tRNA nucleotidyl transferase¹⁴) is suitable; ^3H -borohydride labelled RNA¹² is more generally applicable and we expect this also to be a suitable substrate. The flexibility of using either 5' or 3'

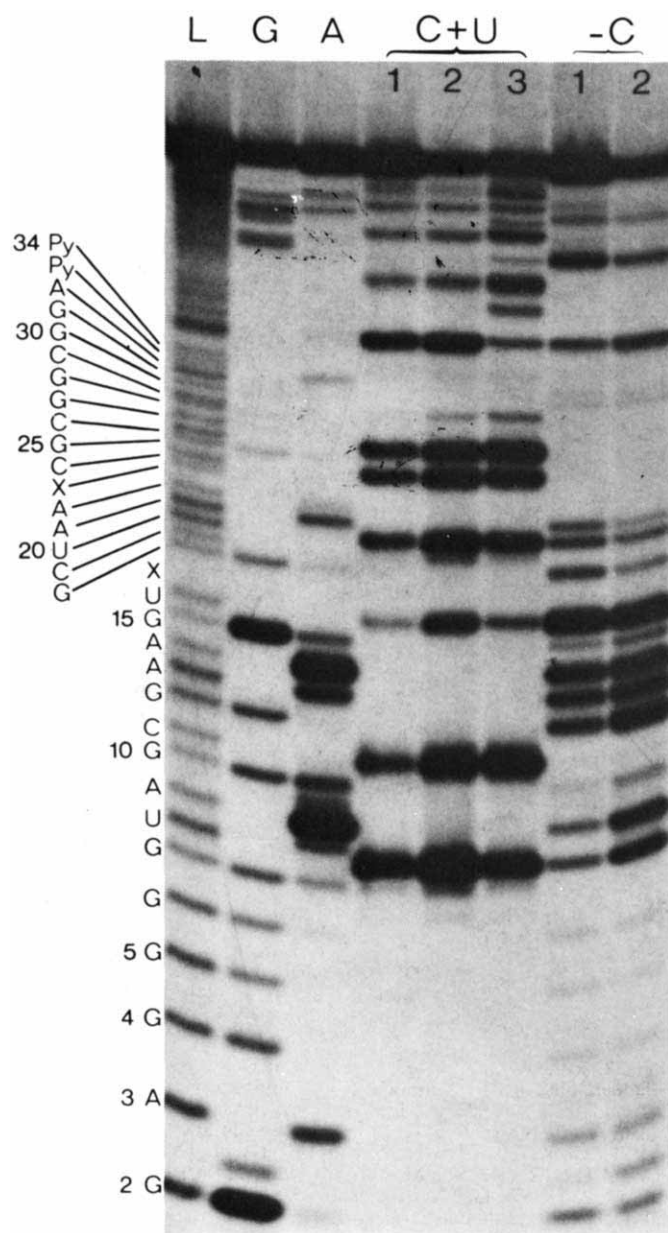


Fig. 3 Sequence analysis of intact tRNA. Letter symbols are as described in Fig. 2. Three separate experiments with pancreatic RNase (C+U, (1) 0.3 ng, (2) 0.5 ng enzyme as in Fig. 2, (3) 0.5 ng enzyme in 70% aqueous buffered formamide, 37°C , 30 min); two with *Phy* I RNase ($1 \mu\text{l}$ enzyme solution, for 5 min and for 20 min) were performed.

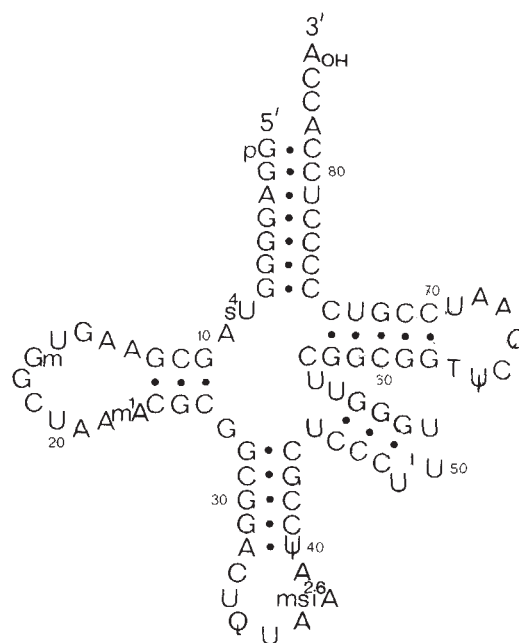


Fig. 4 Sequence of tRNA^{Tyr} of *B. stearothermophilus*.

labelled material should allow the sequencing method to be extended to larger RNA molecules, such as 5S RNA and larger, although there is still a limitation of about 150–200 residues² in the resolution so far obtained in sequencing gels. Finally the method has the potential (not explored here) of probing for conformation of RNA in solution in physiological conditions. Further, the band pattern characteristic of the conformation of the tRNA is likely to be altered when the molecule is complexed with its specific amino acid synthetase. This should give important insight into the specificity of recognition in this tRNA-protein interaction.

While this work was in progress Gupta and Randerath¹² reported a related method using thin-layer chromatography on PEI-cellulose to 'read off' sequences of up to 20 residues in length. As *Phy* I RNase was not used, however, the method was more cumbersome and relied on two-dimensional methods to distinguish the pyrimidine residues. After submission of this report a gel mapping procedure for guanine and adenine residues was reported¹⁵, but a complete sequence could not be established as no method was presented for distinguishing pyrimidine residues.

We thank Professor W. Fiers for suggesting the use of *Phy* I RNase, and Dr J.-P. Bargetzi for providing the enzyme. A.S. is a visiting UNESCO Scholar from the Institute of Biophysics, Hungarian Academy of Sciences, Szeged.

A. SIMONCSITS
G. G. BROWNLEE
R. S. BROWN
J. R. RUBIN
H. GUILLEY

*MRC Laboratory of Molecular Biology,
Hills Road,
Cambridge, UK*

Received 26 July; accepted 2 September 1977.

1. Sanger, F. & Coulson, A. R. *J. molec. Biol.* **94**, 441–448 (1975).
2. Brownlee, G. G. & Cartwright, F. M. *J. molec. Biol.* **114**, 93–117 (1977).
3. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
4. Irwin, M. J., Nyborg, J., Reid, B. R. & Blow, D. M. *J. molec. Biol.* **105**, 577–586 (1976).
5. Brownlee, G. G. *Laboratory Techniques in Biochemistry and Molecular Biology* 13 (North-Holland, Amsterdam, 1972).
6. Schmidt, M., Pipy, D., Wolny, M. & Bargetzi, J.-P. *Experientia* **28**, 739–740 (1972).
7. Pilly, D., Niemeyer, A., Schmidt, M. & Bargetzi, J.-P. *J. biol. Chem.* (in the press).
8. Nishimura, S. *Prog. Nucl. Acid Res. molec. Biol.* **12**, 49–85 (1972).
9. Contreras, R. Ph.D. thesis, Univ. Ghent (1974).
10. Uchida, T., Arima, T. & Egami, F. *J. Biochem. (Tokyo)* **67**, 91–102 (1970).
11. Schmukler, M., Jewett, P. B. & Levy, C. C. *J. biol. Chem.* **250**, 2206–2212 (1975).
12. Gupta, R. C. & Randerath, K. *Nucleic Acids Res.* **4**, 1957–1978 (1977).
13. Contreras, R. & Fiers, W. *Analyst. Biochem.* **67**, 319–326 (1975).
14. Sternbach, H., von d. Haar, F., Schlimme, E., Gaertner, F. & Cramer, F. *Eur. J. Biochem.* **22**, 166–172 (1971).
15. Donis-Keller, H., Maxam, A. M. & Gilbert, W. *Nucleic Acids Res.* **4**, 2527–2538 (1977).

reviews

State of medicine in the US

W. W. Holland

Doing Better and Feeling Worse. Edited by J. H. Knowles. Pp. 287. (W. W. Norton: New York, 1977.) Hardback \$9.95; paperback \$3.95.

THIS interesting collection of essays written by leading personalities in the fields of medicine, politics, economics and the social sciences presents an analysis of opinion of the state of medicine in the United States. As such, it rarely refers to work in other countries but in discussion of such aspects as health insurance, the place of private medicine and development of training for primary care it will be of interest to all those concerned with development of health policies in a number of countries.

The book was written following an invitation by Nelson Rockefeller in his capacity as Vice-President and Chairman of the Commission on Critical Choices for Americans to John Knowles, President of the Rockefeller Foundation, to organise a group to study the problems of health care. A number of the contributions were first submitted as drafts prepared for study by the Commission.

The book covers a wide range from the apparent increase in "medicalisation" of American society, ethical factors such as the right to die—in this particular chapter, "Therapeutic Choice and Moral Doubt in a Technological Age", a detailed description of the Karen Quinlan situation is given—to the place of science and technology, the responsibility of the individual, the challenge of primary care, a look at political aspects of health policy (including relationships between government), the consumers and the providers, and the role of the physician. The book also includes chapters covering financing, research, the needs of children and of mental health.

The major theme of several of these essays does not vary much from that propounded by Thomas McKeown in his recent text *The Role of Medicine* or that of Ivan Illich in *Medical Nemesis: The Expropriation of Health*. Aaron Wildavsky, author of the chapter entitled "Doing Better and Feeling Worse: The Political Pathology of Health Policy", states that the great equation Medical Care equals Health is wrong. "More available medical care

does not equal better health. The best estimates are that the medical system (doctors, drugs, hospitals) affects about 10% of the usual indices for measuring health: whether you live at all (infant mortality), how well you live (days lost due to sickness), how long you live (adult mortality)". He continues by saying that since the other 90% are affected by life-style and social and environmental conditions, "most of the bad things that happen to people are at present beyond the reach of medicine".

Walsh McDermott, in his chapter entitled "Evaluating the Physician and His Technology", takes this point a bit further by saying that it is wrong to say that the personal encounter physician has little influence on health but he lacks influence on the indicators of community health largely devised in the eighteenth, nineteenth and twentieth centuries for a different purpose. He continues, however, by saying that we in fact lack measures of the physician's ability to "modulate the deranged physiology of ultimately fatal chronic disease". McDermott's conclusion on the note that doctors play, coincides with that of several other authors.

Another valuable point made by several authors is that a half-way stage has been reached in the development

of medicine. Major advances have been made but the gains in the past have been fairly minimal compared with what might be expected in future. The book contains the underlying messages of the need to change life-style and the false expectations that populations have of medicine. Further points particularly mentioned by David Callahan in his chapter, "Health and Society: Some Ethical Imperatives", are "the breakdown of the ethical distinction between need and desire in our culture" and also "the continuing utopian lure of technology, a lure whose net effect is to thwart any attempt to place limits on medical aspirations".

This book, despite its impressive variety of subject matter, does not really state much that is new. The particular question, however, that comes to mind is can such publications as this and that of McKeown, Cochrane and Illich have any influence on consumers, providers, or governments; and if they can, will this result in greater rationalisation of requirements for health care? □

W. W. Holland is Professor in the Department of Community Medicine at St Thomas's Hospital Medical School, London, UK.

Control of cellular differentiation

The Differentiation of Cells. By N. Maclean. Pp. iii+216. (Edward Arnold: London, 1977.) Hardback £12; paperback £5.95.

THIS volume, the first of a series entitled *Genetics-Principles and Perspectives*, covers one of the central problems in developmental biology—the control of cell differentiation. There are now several recent texts which cover aspects of the control of gene expression in developing systems. The present volume presents a wide ranging view of current problems.

The first chapter discusses multicellularity in relation to the evolutionary significance of differentiation. Thereafter, the chapters cover differential gene expression, control mechanisms in gene expression, the cytoplasm in differentiation, the role of

hormones, episomes and viruses, the cell surface, and cancer and differentiation. The book ends with the author's views of the important unanswered questions in this field and the experimental systems that hold special promise for their elucidation.

The wide range of subject matter reflects the enormous range of controls which may operate in the orderly progression of cells to their terminally differentiated states, and the maintenance of these states once achieved. A major aspect of the control of cell differentiation is the regulation of gene activity. In introducing the problem of differentiation Dr Maclean emphasises most strongly that "differential gene activity may be a result of differentiation, rather than its causal mechanism". I found this introduction confusing because the terminology is not defined. It is unfortunate that the

book does not have a glossary, since, as the author points out in the case of heterochromatin, the implication of particular terms can change with time.

The subject material is laid out in a readily accessible manner and profusely illustrated with diagrams. The significance of some of the illustrations to the text (for example, the fibre autoradiographs p. 94) is, however, not clearly stated, and in the section on histones only the old nomenclatures are used. At various places the clarity of information has been enhanced by tabulation, though this format in some cases (for example, observations explicable by the master-slave hypothesis) has limited discussion of points which

may have alternative explanations.

A wide range of experimental systems and approaches have been used throughout the text. Perhaps as a consequence of this there are places where advanced students will need greater detail. The questioning style in which the text is presented should motivate the undergraduates and graduates for whom it was written to follow up the references to specialist reviews, and stimulate them to deeper consideration of the complex but intriguing problems of cell differentiation.

C. C. Ford

C. C. Ford is Lecturer in Developmental Biology at the University of Sussex, UK.

Lunar science

The Moon Book: Exploring the Mysteries of the Lunar World. By Bevan M. French. Pp. 287. (Penguin: New York and Harmondsworth, UK, 1977.) \$4.95; £1.95.

THIS is a book for the layman, as might be guessed from the name of the publisher rather than from the presumptuous title. A reader who has some general acquaintance with the Earth sciences, however, will find considerably greater profit in it than the 'intelligent layman' who starts from a state of near-total ignorance.

May natural inclination was to turn first to chapter 10, entitled, "The Lunar Interior: Mapping with Echoes". It contains a disappointing discussion of deductions made from geophysical evidence about the lunar interior. The seismic evidence which has been included seems to be a simplified rehash of the appropriate section of Ross Taylor's 1975 book, referenced in the bibliography.

There are some elementary errors, such as the statement that the inner core of the Earth is under a pressure of nearly a million atmospheres—in fact, it is over three million atmospheres—and a confusion between kilometres and miles in the figure showing cross-sections of the Earth and Moon. But the worst error I found was in the discussion about the Moon's ancient magnetic field, and whether it could have been provided by a molten iron core which has since frozen. French suggests this is unlikely because, "A large body like the Moon loses heat only by the slow conduction of heat through its interior." Ignorance of modern ideas about convection in planets, with the consequent ability for planetary interiors to lose heat at rates several orders of magnitude above conduction

rates, is appalling in a work of this kind.

The author also seems to confuse the ideas of a lunar asthenosphere and a possibly molten lunar core. Deep moonquakes occur in a zone between 600 and 1,000 km from the surface, whereas the lunar radius is 1,738 km. Seismic signals penetrating below this depth have rarely been observed. This suggests attenuation in a region where the pressure is similar to that in the Earth's low velocity zone, at a depth of around 100 km; so, by analogy it may be inferred that the lunar mantle at 1,000 km depth must be in a similar state to the material in the Earth's low velocity zone, possibly including some partial melt, in order to cause the seismic attenuation. Best estimates for the Moon's moment of inertia factor, plausibility arguments about planetary heat transfer and another tenuous piece of seismic evidence, not available to Ross Taylor, combine to suggest that if the Moon does have a metallic iron core, it is probably solid and not more than 500 km in radius.

In the rest of the book the author is on more familiar ground. He gives a very readable account of the pre-Apollo ideas of the Moon's origin and history, the Apollo missions themselves, and the knowledge gained from studies of the returned samples of rock and lunar soil. A set of thirty-six black and white photographs illustrate the story well. The final chapter rounds off the book nicely with a discussion of future developments.

I can recommend this book for general reading provided that the more technical reader bears in mind the author's qualifications as a geologist and geochemist. It would benefit substantially from a rewriting of chapter 10, possibly by a co-author.

Neil R. Goulty

Neil R. Goulty is Research Associate at the Department of Geophysics and Planetary Physics, University of Newcastle-upon-Tyne, UK.

Plasma low density lipoproteins

Low Density Lipoproteins. Edited by C. E. Day and R. S. Levy. Pp. xx+445. (Plenum: New York and London, 1977). \$47.40.

THIS book, written by several authors, provides up-to-date information on the structure and metabolism of plasma low-density lipoproteins (LDL), together with an examination of their comparative biology and their relationship to disorders of lipid metabolism and atherosclerosis. The material is grouped in five sections, comprising a total of 15 chapters which provide good critical coverage of the various topics.

The reader is likely to appreciate the efforts of the editors, authors, and publisher to condense a large body of information into a single, well-organised volume. For example, the section devoted to the discussion of "Aberrations of Metabolism" presents a rather detailed description of the pathogenesis and treatment of a few selected genetic disorders, identified as hyper- β and hyperpre- β lipoproteinaemia, type III hyperlipoproteinaemia (also known as "broad β disease" or dysbetalipoproteinaemia), and abetalipoproteinaemia, together with an analysis of the variants of LDL with particular reference to the Lp(a) antigen.

The section on "Comparative Biology", which gives a good overview on the LDL of mammals and non-mammalian vertebrates, is of particular interest. Concerning atherosclerosis, a useful discussion is presented on the interactions *in vitro* of LDL with either small polyanions (heparin and dextran sulphate) or macromolecules (collagen and elastin); these interactions are believed to play a role in the deposition of LDL in the arterial wall and in atherogenesis.

Perhaps the least rewarding section for both authors and readers may prove to be that dealing with LDL structure, since much of the current information, particularly on the LDL protein, is rather uncertain and is still the subject of an active research pursuit.

Overall, this is a good book, especially for investigators in the lipoprotein field who are likely to find both the condensed information and the more than 1,500 citations very valuable for their work.

Angelo M. Scanu

Angelo M. Scanu is Professor in the Departments of Medicine and Biochemistry, University of Chicago, Pritzker School of Medicine, Chicago, Illinois.

Applying physics to chemistry

Modern Physics in Chemistry. Vol. 1. Edited by E. Fluck and V. I. Goldanskii. Pp.xiii+406. (Academic: New York, San Francisco and London, 1977.) £18.50; \$36.10.

THIS is a collection of papers intended to describe the basic principles of some of the experiments of modern physics which have in the past decade been applied with such signal success to the problems of chemistry. It is not, as could easily have happened, primarily a collection of the various new spectroscopies, although inevitably resonance methods feature strongly. Chapters on positronium and mesic chemistry vary the diet.

This is not a unified, critical or comparative text which would serve, say for the education of physical chemists in newer methods. This is a disappointment for there is always room for well written interdisciplinary teaching books of this kind. It may, however, be interesting for the researcher to have under one cover such a collection of reviews which, though they are now rather dated (the manuscripts were collected in 1973), provide a useful insight into the Russian literature in some of the areas.

Two chapters on X-ray spectroscopy divide into valence shell and band structure and core level shifts, respectively. They are, however, really too brief to provide a useful

survey of this rather slowly but long developing field. Electron paramagnetic and nuclear quadrupole resonance methods have much fuller coverage. Nuclear magnetic resonance is not featured; it is not "modern" any longer, except in the form of chemically induced dynamic nuclear polarisation. Mossbauer spectroscopy is brought up-to-date with Mossbauer double resonance experiments.

Although a different viewpoint on these newer techniques is to be welcomed, it is rather difficult to see much benefit being derived from this first volume by anyone outside the ranks of the historians of science. Unfortunately, in translation, clarity has somehow escaped and even post-graduate students will find quite a lot of up-hill work, sometimes intelligent guesses being necessary to perceive exactly what the authors mean. There is little evidence of editorial effort to attain a unity of style or to secure an error-free translation (Gammett constant?) although stilted or archaic usage can, as the late Gerard Hofnung showed, have charm. More serious slips are to be found, including a blatantly erroneous diagram of nodal surfaces for the benzene π orbitals. It seems as if the chemists' voice has been too weak throughout. It is difficult to see who would be recommended to pay £18.50 for this volume.

D. W. Turner

D. W. Turner is Lecturer in Physical Chemistry at the University of Oxford, UK.

Gas chromatography methodology

Modern Practice of Gas Chromatography. Edited by R. L. Grob. Pp. (Wiley-Interscience: New York and London, 1977.) £16; \$28.

THERE are now so many good books on the theory and practice of gas chromatography that a newcomer to the ranks must be of high quality indeed to stand any chance of acceptance. The book under review has the now almost universal style of a multiplicity of authors, each of whom holds responsibility for his own chapter and the editor for the resultant patchwork. The subjects covered separately are: Introduction (history and nomenclature); Theory and Basics; Columns and Column Selection; Qualitative and Quantitative Analysis; Detectors; Instrumentation; Trace Analysis; Selection of Analytical data (data handling); Analysis of Foods; Clinical Applications; Physico-Chemical Measurements;

and Drug Analysis.

The sections are very variable. The best is by Touchstone and Dobbins on "Clinical Applications", where the descriptions are clear, concise and accurate and the authors have actually read papers appearing before 1968. One of the less good sections is that dealing with instrumentation where the illustrations are frequently commercial black boxes that convey no information at all (except advertising). The editor also needs to note that the same illustration is used on page 355 as on page 173 in a different chapter.

All-in-all, the book is rather like the curates egg (good in parts); and with a well understood and practised technique like gas chromatography, this is not good enough. This is a pity, since a lot of hard work has gone into the chapters; a clearer intent as to what needs to be put across in each section would have produced a shorter and better book.

A. T. James

A. T. James works in the Colworth Laboratory of Unilever Ltd, Sharnbrook, UK.

Organosulphur chemistry

Organic Chemistry of Sulfur. Edited by S. Oae. Pp. ix+713. (Plenum: New York and London, 1977.) \$48.

THE importance of organosulphur chemistry, as demonstrated by its rapid growth over the past decade or so, has meant that a book covering this area in breadth and depth is somewhat overdue. Twelve chapters on various aspects are contained in the book and these and their contents are arranged in a logical sequence. Each of the chapters is written by one or more leading authorities; and it is a compliment to the contributors and to the editor that, although individual styles are apparent, continuity is maintained.

In the first chapter, the basic theoretical aspects of sulphur bonding are described. This is followed by a discussion of elemental sulphur and its reactions, which provides a necessary introduction to the vulcanisation of rubber covered in the third chapter. The remaining chapters deal with the physical and chemical properties of specific sulphur compounds—thiols, thiones, sulphides (mono-, di-, and poly-), sulfoxides and sulphilmines, sulphonium salts, sulphones and sulfoximines, sulphonic acids and esters, and sulphonate and sulphate esters. Some detail is also given to sulphur biochemistry within various chapters.

On the credit side, the book contains many excellent qualities with the diagrams, tables and general layout being attractive and comprehensive. The mechanistic approach used is also a highly commendable feature.

On the debit side, however, it is surprising that the industrially important sulphonic acids are not discussed in some detail. Furthermore, over 2,300 references are quoted, but no references later than 1973 are included and very few 1972 or 1973 references are given. In fact, for the majority of the chapters, the literature is covered adequately only to the end of 1971 and for one chapter only to the end of 1970. This overlong delay between completion of the manuscript and publication must detract in some degree from the overall usefulness of the book.

In spite of these drawbacks, and the presence of a considerable number of trivial and typographical errors, the book is worthy of a place in chemistry libraries and on the bookshelves of the more affluent workers in the field.

A. E. Tipping

A. E. Tipping is Reader in Organic Chemistry at the University of Manchester Institute of Science and Technology, Manchester, UK.

Karst landforms

Morphogenetics of Karst Regions. By L. Jakucs. Pp. 284. (Adam Hilger: Bristol, UK, 1977.) £17.

THIS book is a revised and enlarged version of an original text by the Hun-

garian geographer, Professor Jakucs. The generally good translation is by B. Balkay. In recent years there has been a number of books in different European languages on the landforms of karst (massive limestone) regions, but this book is the first to focus primarily on the processes and origins of karst landforms. It therefore not only fills a gap in our literature on karst but is

also a highly perceptive contribution to the subject.

After discussing the concepts of karst and of karst corrosion Jakucs treats his subject under the main factors which control karstification. These he lists as petrovance; epirogenic movements (that is, structural variance); climatic variance; erosional and geomorphological variance; and anthropovariance. Each of these controls is discussed at length, and this forms the main theme of the book. The author aims to show that all variations of karst can be discussed in the light of the variants proposed.

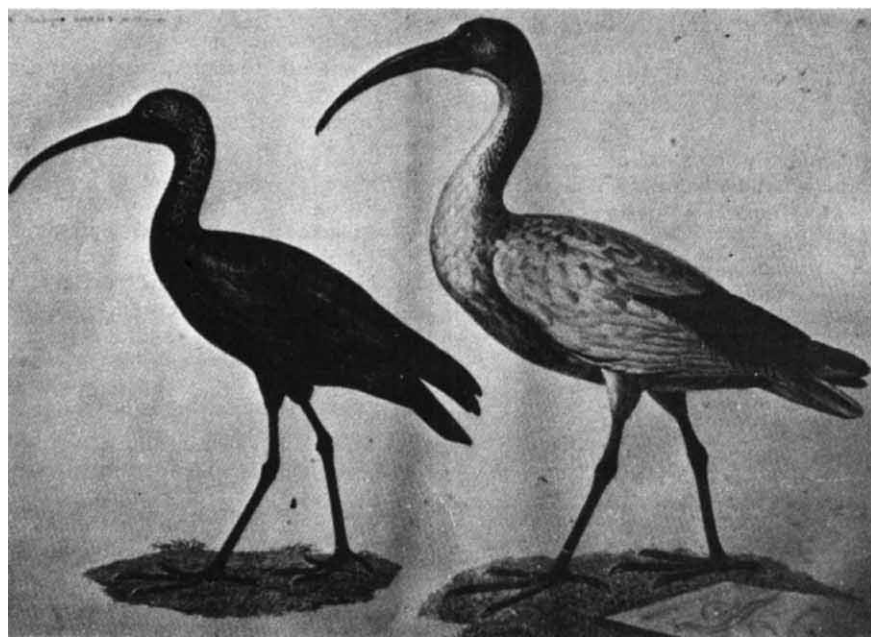
The treatment of each section is comprehensive and the discussion covers some ground familiar to karst geomorphologists. But the text is illuminated by Jakucs' own illustrations and careful experiments, which form the main scientific contribution of the book. There are certain sections which are therefore new and also controversial to the English karst morphologist. Four aspects are of particular interest. The first is that dealing with the effects of limestone crystallinity and lithostructure on karst corrosion. The second discusses the effects of soil microclimates on the development of karst depressions. Thirdly, the author has some interesting observations to make on the role of geomorphological situation and erosional variations on the origin of karst landforms; he usefully differentiates between authigenic and allogenic karst evolution. And fourthly there is an extended discussion on the effects of man on karst evolution; the results of deforestation, of spring development and of dam building are among the topics included. The author is perhaps least at home when dealing with the effects of climatic factors on the development of karst; even here, however, he has some shrewd comments to make, though he has had to rely mainly on the published literature.

The book is well illustrated with clear diagrams and photographs. The bibliography is especially good on Central and Eastern European works, but shows a lack of knowledge of the more recent Western European and American work (that is, since 1970).

The Hungarian school of karst morphology has long been distinguished for its contributions to all aspects of karst, both surface and underground. Professor Jakucs' book illustrates the depth and grasp of Hungarian work in this field, and his succinct work should be carefully read by all English workers in karst geomorphology.

M. M. Sweeting

M. M. Sweeting is Lecturer in Geography at the University of Oxford, UK.



Illustrations taken from *Food: The Gift of Osiris* by W. J. Darby, P. Ghalioungui and L. Grivetti (Academic: London, New York and San Francisco, 1977; two volumes £12.50/\$27 and £12.50/\$25). These two volumes detail exactly what is known about ancient Egyptian food and eating habits. The book considers medical as well as nutritional uses, and provides a wealth of archeological and anecdotal information. *Top*, Sacred ibis drawn by scientists accompanying Bonaparte in 1799. The bird was associated with the god Thot, master of knowledge and secret science. Strict measures were taken to protect the bird in ancient Egypt, but it has since become extinct in that region. *Bottom*, Hand fish net (tomb of Ka-Gem-Ni at Saqqara, Old Kingdom, sixth dynasty, reign of King Teti, about 2390 BC). Open mouth V nets were used to collect fish which had been channelled into a fenced barricade, and this type of net is still in use today.